ELSEVIER Stochastic Processes and their Applications 62 (1996) 191–222

stochastic processes and their applications

High-density limits of hierarchically structured branching-diffusing populations

Donald A. Dawson[a,1], Kenneth J. Hochberg[b,*,2], Vladimir Vinogradov[c,3]

[a] *Department of Mathematics and Statistics, Carleton University, Ottawa, Ont. Canada K1S 5B6*
[b] *Department of Mathematics and Computer Science, Bar-Ilan University, 52900 Ramat-Gan, Israel*
[c] *Department of Mathematics and Computer Science, University of Northern British Columbia, Prince George, BC Canada V2N 4Z9*

Received May 1995; revised January 1996

Abstract

We develop a general probabilistic approach that enables one to get sharp estimates for the almost-sure short-term behavior of hierarchically structured branching-diffusion processes. This approach involves the thorough investigation of the cluster structure and the derivation of some probability estimates for the sets of rapidly fluctuating realizations. In addition, our approach leads to the derivation of new modulus-of-continuity-type results for measure-valued processes. In turn, the modulus-of-continuity-type results for hierarchical branching-diffusion processes are used to derive upper estimates for the Hausdorff dimension of support.

Keywords: Hierarchical branching; Path properties; Modulus of continuity; Hausdorff dimension

AMS classification: Primary 60F15, 60G17; 60J60; 60J80

0. Introduction

In this paper, we continue the investigation of path properties of *hierarchically structured measure-valued branching-diffusion processes* that we began in our two earlier works (Dawson, Hochberg and Vinogradov, 1994, 1995, hereafter referred to as [DHV1] and [DHV2], respectively). These processes describe populations of individuals undergoing some spatial motion that are affected by one branching mechanism that acts upon individuals and by another, independent, branching mechanism

* Corresponding author.

[1] Supported by an NSERC grant.
[2] Partially supported by an NSERC International Scientific Exchange Award and a grant from the Bar-Ilan University Research Authority.
[3] Partially supported by an NSERC grant, NSERC grants of D.A. Dawson and J.N.K. Rao and an internal travel grant from the Department of Mathematics and Computer Science of the University of Northern British Columbia.

0304-4149/96/$15.00 © 1996 Elsevier Science B.V. All rights reserved
PII S0304-4149(96)00052-X

that acts simultaneously upon groups of individuals. Such "upper level" groups can be thought of as representing superparticles or families or clans. A critical feature of such hierarchical (or *multilevel*) processes is the absence of independence in the branching, as a result of simultaneous branching at the upper (family or clan) level. This increases the dependencies in the system and, along with the presence of a hierarchical structure, necessitates a finer analysis of the path behavior and influences both the short-term and long-term behavior of the processes in many ways. Such processes were first introduced in Dawson, Hochberg and Wu (1990) and Dawson and Hochberg (1991) and generalize in a natural way nonhierarchical *measure-valued branching-diffusion processes* or *superprocesses*. Our consideration of *hierarchical* measure-valued processes is motivated by various examples from molecular biology, population genetics, computer science, and cultural diffusion. (Some of these are described in [DHV2], as well as in several of the papers quoted in the list of references.) In particular, we refer the reader to Dawson and Wu (1996), where a *dynamical-information-structure model* is considered.

It is worth noting that hierarchical branching-diffusion processes considered in this work emerge as weak (or strong) limits of certain appropriately rescaled *hierarchical branching-diffusing particle systems* (BPS-2). Below, we will describe such hierarchical branching-diffusing particle systems and also characterize their high-density limits. In Section 2, we formulate the main results of this paper. In Section 3, we discuss a characterization of superprocesses via Laplace functionals.

Let us emphasize that our approach turns out to be appropriate for the derivation of modulus-of-continuity-type results, not only for hierarchical measure-valued processes, but even for the classical Wiener process (see Theorem 2.1(i), which provides the exact almost-sure rate of convergence in P. Lévy's global modulus of continuity for d-dimensional Wiener process). The complete proof of Theorem 2.1(i) can be found in Dawson et al. (1966a), hereafter referred to as [DHV3].

In Section 4, we first describe the cluster structure of hierarchical branching-diffusion processes and then present the almost-sure lower estimate for the local propagation of mass from a point source. The approach used in that section develops along the same lines as that pursued in Dawson and Vinogradov (1994), hereafter referred to as [DV], in the single-level setting. Note that we only state and briefly describe the major ideas of the proof of Theorem 4.4 (the main result of that section), owing to its similarity with the case of *non-hierarchical* branching. Let us point out that the approach developed in Section 4 is more comprehensive than the description of the cluster structure of hierarchical branching-diffusion processes presented in [DHV1]. Note that in order to prove Theorem 4.4, one should employ the cluster structure of superprocesses along with arguments related to the strong convergence of hierarchical branching-diffusing particle systems to super-2 processes. In turn, the strong convergence of such hierarchical systems can be established by the use of arguments similar to those used in Mueller and Perkins (1992) and Barlow and Perkins (1994) in the single-level setting.

Subsequently, we derive a number of almost-sure upper estimates in Section 5 and then apply some of these results in Section 6 to obtain an almost-sure upper estimate for the Hausdorff–Besicovitch dimension of the closed support of the aggregated process.

It should be mentioned that in this paper we do not consider at all the interconnections between the theory of weak convergence of branching-diffusing particle systems and the classical theory of summation of independent random variables that were sketched in Subsection 1.2 of [DHV2] (cf., e.g., Proposition 1.2.1 therein). Such interconnections are of independent interest and are considered in detail in Dawson et al. (1996b). Also, note that hierarchically structured branching-diffusion measure-valued processes and their various properties have recently been studied in several other works quoted in the list of references.

1. Heuristic description of the model and the branching particle picture

Note that the main results of this paper deal with sample path properties of multilevel superprocesses. However, both the construction of these processes and the proofs of the sample path properties are based on the use of the corresponding branching-diffusing particle systems. Therefore, before stating our main results, we will give in this section a detailed description of the approximating branching particle systems. We note that certain points of the following description were given to some extent in [DHV1] and [DHV2]. Here, we present a detailed description of a slightly simplified version of the model, in order to make the present paper easier to read.

First, we sketch briefly the main features of single-level branching models; we will then impose a hierarchical structure with the introduction of a second level of branching. (We refer to Dawson (1993), hereinafter referred to as [D], for a systematic consideration of such single-level structures.) To this end, we introduce the following two empirical processes, which play a very important role in the theory of branching-diffusion processes:

$$U_t^{(\eta)} := \frac{1}{\eta} \sum_{j=1}^{N(t)} \delta_{x_j(t)} \quad \text{and} \quad H_t^{(\eta)} := \frac{1}{\eta} \sum_{j=1}^{N(t)} \delta_{x_j(\cdot \wedge t)}.$$

Here, $N(t)$ denotes the random (because of the branching) number of particles at time t, $x_j(t)$ are their sites, and $x_j(\cdot \wedge t)$ is a trajectory of the jth particle stopped at time t. Note that $U_t^{(\eta)} \in M_{\mathrm{F}}(\mathbb{R}^d)$. If the trajectories $x_j(\cdot)$ are assumed to be continuous, then $H_t^{(\eta)} \in M_{\mathrm{F}}(\mathbb{C}_d^{(0)})$, where

$$\mathbb{C}_d^{(0)} := \mathbb{C}([0, \infty), \mathbb{R}^d).$$

The probability law of the process $\{H_t^{(\eta)}\}_{t \geqslant 0}$ then belongs to $M_1(\mathbb{D}_d^{(1)})$, where

$$\mathbb{D}_d^{(1)} := \mathbb{D}([0, \infty), M_{\mathrm{F}}(\mathbb{C}_d^{(0)})),$$

and where $M_{\mathrm{F}}(E)$ and $M_1(E)$ denote, respectively, the spaces of finite measures and of probability measures on the space E, furnished with the topology of weak convergence.

The (non-historical) empirical process $U_t^{(\eta)}$ is obtained from the historical empirical process $H_t^{(\eta)}$ by projection. In order to describe this rigorously, we now introduce the following two projection operators:

$$\pi_t(\omega) := \omega(t) \quad \text{and} \quad \tilde{\pi}_t(\mu)(A) := \mu(\{y\colon \ y_t \in A\}),$$

where $\mu \in M_{\mathrm{F}}(\mathbb{C}_d^{(0)})$, and A belongs to the σ-algebra $\mathscr{B}(\mathbb{R}^d)$ of Borel sets on $\mathbb{R}^d$. It then clearly follows that

$$U_t^{(\eta)} := \tilde{\pi}_t(H_t^{(\eta)}). \tag{1.1}$$

Next, we describe the branching mechanism. Assume that we start with a Poisson number $\Pi(\eta)$ of independent particles located at the origin. Each particle is assigned mass $1/\eta$ and is assumed to perform an independent spatial motion. (In this paper, we consider only the case in which the spatial motions of all particles are d-dimensional Wiener processes taking values in $\mathbb{R}^d$.) We also assume that at an exponentially distributed instant of time with mean $\mu(\eta) = \eta^{-\beta}$, the particle located at site x splits into a random number of offspring that can depend on the site x (*local branching*). For the sake of simplicity, we assume that the mechanism of the local branching is critical, space homogeneous, and governed by the particle-production generating function

$$\vartheta_x^{(\eta)}(u) := \vartheta(u;\beta) = u + \frac{1}{1+\beta}(1-u)^{1+\beta}, \tag{1.2}$$

where $\beta \in (0,1]$ is fixed.

Remark 1.1. Note that in fact the constant $1/(1+\beta)$ from (1.2) can be replaced by an arbitrary positive constant. The case of an arbitrary positive constant is reduced to our special case by scaling in space and time.

It is well known that the above-described branching-diffusing particle system $U_t^{(\eta)}$ converges weakly to a càdlàg $M_{\mathrm{F}}(\mathbb{R}^d)$-valued process U_t as $\eta \to \infty$ (cf. Section 4 of [D]) and that $H_t^{(\eta)}$ converges weakly to a càdlàg $M_{\mathrm{F}}(\mathbb{C}_d^{(0)})$-valued process H_t (cf. Dawson and Perkins, 1991) as $\eta \to \infty$. Hereafter, we refer to U_t as the $(2,d,\beta)$-*super*-1 *process*. This limiting super-1 process is uniquely characterized by its Laplace functional (see Section 3 for more details), which in turn is characterized by the motion mechanism (d-dimensional Weiner process) and the function

$$\Phi_x(\lambda) = -\frac{1}{1+\beta}\lambda^{1+\beta}$$

which is closely related to the particle-production generating function $\vartheta(u;\beta)$ given by (1.2).

In addition, the limit as $\eta \to \infty$ of the *historical* version $H_t^{(\eta)}$ of the branching-diffusing particle system $U_t^{(\eta)}$ is hereafter referred to as the $(2,d,\beta)$-*historical process* H_t. Note that H_t is a càdlàg $M_{\mathrm{F}}(\mathbb{C}_d^{(0)})$-valued process if $\beta < 1$ and a continuous $M_{\mathrm{F}}(\mathbb{C}_d^{(0)})$-valued process if $\beta = 1$ (cf. Dawson and Perkins, 1991). In particular, the latter implies that the case $\beta = 1$ leads to the subset of $\mathbb{D}_d^{(1)}$ which consists of continuous sample paths (cf., e.g., Theorem 7 of El Karoui and Roelly (1991) or Theorem 6.1.3 of [D]). However, in this paper we have chosen to use the notation $\mathbb{D}_d^{(1)}$ in both cases to allow for a unified treatment. Note also that we can take limits as $\eta \to \infty$ on both sides of (1.1) to obtain

$$U_t := \tilde{\pi}_t(H_t). \tag{1.1'}$$

Now, let $\mathbf{P}_\mu \in M_1(\mathbb{D}([0,\infty), M_{\mathrm{F}}(\mathbb{R}^d)))$ denote the probability law of the super-1 process U_t, where $\mu \in M_{\mathrm{F}}(\mathbb{R}^d)$ denotes the initial measure. Let $\mathbb{Q}_{0,\mu} \in M_1(\mathbb{D}_d^{(1)})$ denote the probability law of the corresponding historical process H_t, where the subscript μ is to be understood as before.

We now proceed to the description of *hierarchical* processes. To this end, consider a system of diffusing particles in $\mathbb{R}^d$ with hierarchical structure. Namely, assume that a Poisson number $\Pi(\eta_2)$ of independent families of particles (hereafter referred to as *level-2 particles* or *superparticles*) is given. Mass $1/\eta_2$ is assigned to each superparticle, and the ith superparticle is assumed to consist of a Poisson number $\Pi_i(\eta_1)$ of independent particles (*level-1 particles*). Here, all the Poisson numbers $\Pi(\eta_2)$ and $\Pi_i(\eta_1)$, $0 \leqslant i \leqslant \Pi(\eta_2)$, are assumed to be independent of each other and of everything else. Ultimately, we will let η_1 and η_2 approach infinity. We assume for simplicity that all the particles are located at the origin at time $t = 0$ and perform independent Brownian motions in $\mathbb{R}^d$. We now assign mass $1/\eta_1$ to each level-1 particle. It is assumed that any individual level-1 particle splits into a random number of offspring at an exponentially distributed instant of time with mean $\eta_1^{-\beta_1}$, and that each newly born particle is a copy of its parent and immediately starts to perform d-dimensional Brownian motion. The motions, lifetimes and branchings of all particles, as well as the initial number of superparticles and the initial number of particles within each initial superparticle, are independent of each other and of everything else. Each superparticle is also assumed to split (independent of everything else) into a random number of superparticles, each of which copies its parent superparticle. It is natural to assume that the superparticle-lifetime distribution function is exponential with mean $\eta_2^{-\beta_2}$.

We next give a rigorous description of the mechanism of hierarchical branching. By analogy with the case of single-level branching, assume that the mechanism of the local branching in the first hierarchical level is governed by the particle-production generating function given by (1.2) with $\beta = \beta_1 \in (0,1]$. The mechanism of the local branching on the second hierarchical level is assumed to be governed by the superparticle-production generating function given by (1.2) with $\beta = \beta_2 \in (0,1]$.

Now, we present a rigorous description of the BPS-2, introduced above on a heuristic level, via two empirical processes:

$$X_t^{(\eta_1,\eta_2)} := \frac{1}{\eta_2} \sum_{i=1}^{N_2(t)} \delta_{\left\{\frac{1}{\eta_1}\sum_{j=1}^{N(i,t)} \delta_{x_{i,j}(t)}\right\}}$$

and

$$\mathscr{H}_t^{(\eta_1,\eta_2)} := \frac{1}{\eta_2} \sum_{i=1}^{N_2(t)} \delta_{\left\{\frac{1}{\eta_1}\sum_{j=1}^{N(i,s\wedge t)} \delta_{x_{i,j}(\cdot\wedge s\wedge t)}\right\}_{s\geqslant 0}}$$

where $N_2(t)$ denotes the number of superparticles alive at time t and $N(i,t)$ denotes the number of particles in the ith superparticle at time t. Also, summation over i denotes summation over all superparticles, summation over j denotes summation over all level-1 particles belonging to the ith superparticle, $x_{i,j}(t) \in \mathbb{R}^d$ denotes the location at time t of the jth level-1 particle belonging to the ith superparticle, and $x_{i,j}(\cdot \wedge t)$ denotes a trajectory of the jth level-1 particle belonging to the ith superparticle, stopped at time t. Note that at each time instant t, the empirical process $X_t^{(\eta_1,\eta_2)}$ is an $M_{\mathrm{F}}(M_{\mathrm{F}}(\mathbb{R}^d))$-valued process, since each individual superparticle can be viewed as an $M_{\mathrm{F}}(\mathbb{R}^d)$-valued process. Therefore, the configuration of superparticles represents a finite measure on the space $M_{\mathrm{F}}(\mathbb{R}^d)$, i.e., it is indeed an element of $M_{\mathrm{F}}(M_{\mathrm{F}}(\mathbb{R}^d))$. Also, $\mathscr{H}_t^{(\eta_1,\eta_2)}$ is a càdlàg $M_{\mathrm{F}}(\mathbb{D}_d^{(1)})$-valued process, that is, its sample paths lie in

$$\mathbb{D}_d^{(2)} := \mathbb{D}([0,\infty), M_{\mathrm{F}}(\mathbb{D}_d^{(1)})).$$

Note that by analogy with (1.1) and (1.1′), on can obtain the following relationships between $X_t^{(\eta_1,\eta_2)}$ and $\mathscr{H}_t^{(\eta_1,\eta_2)}$ as well as between their corresponding limits (as $\eta_1 \to \infty$ and $\eta_2 \to \infty$) X_t and $\mathscr{H}_t$:

$$X_t^{(\eta_1,\eta_2)} = \hat{\pi}_t(\mathscr{H}_t^{(\eta_1,\eta_2)}), \tag{1.3}$$

and

$$X_t = \hat{\pi}_t(\mathscr{H}_t). \tag{1.3′}$$

Here, the projection operator

$$\hat{\pi}_t\, M_{\mathrm{F}}(\mathbb{D}_d^{(1)})) \to M_{\mathrm{F}}(M_{\mathrm{F}}(\mathbb{R}^d))$$

is defined by means of the projection operator $\tilde{\pi}_t$ introduced via

$$\hat{\pi}_t(\nu(B)) := \nu(\mu: \tilde{\pi}_t(\mu_t) \in B), \tag{1.4}$$

where B belongs to the σ-algebra $\mathscr{B}(M_{\mathrm{F}}(\mathbb{R}^d))$ of Borel sets on $M_{\mathrm{F}}(\mathbb{R}^d)$, the measure μ_t belongs to $M_{\mathrm{F}}(\mathbb{C}_d^{(0)})$, and the measure ν belongs to $M_{\mathrm{F}}(\mathbb{D}_d^{(1)})$.

Recall that it often suffices to consider only *level*-1 *projections* of the empirical process $X_t^{(\eta_1,\eta_2)}$ (see, e.g., the dynamical-information-structure model described above). In particular, this might be useful if our interest centers mainly on configurations of level-1 particles, in which case it is not too important to know to which level-2 particles those level-1 particles belong. In this respect, we now rigorously introduce the (*non-historical*) *aggregated process* $Z_t^{(\eta_1,\eta_2)}$ and the *historical aggregated process* $\mathscr{Z}_t^{(\eta_1,\eta_2)}$ that correspond to the empirical processes $X_t^{(\eta_1,\eta_2)}$ and $\mathscr{H}_t^{(\eta_1,\eta_2)}$ introduced above, namely,

$$Z_t^{(\eta_1,\eta_2)} := \frac{1}{\eta_1\eta_2}\sum_{i,j} \delta_{x_{i,j}(t)} \tag{1.5}$$

and

$$\mathscr{Z}_t^{(\eta_1,\eta_2)} := \frac{1}{\eta_1\eta_2}\sum_{i,j} \delta_{x_{i,j}(\cdot\wedge t)}. \tag{1.5′}$$

Recall that the heuristic meaning of the aggregated process was explained at the beginning of this section. Clearly, at each time instant t, $Z_t^{(\eta_1,\eta_2)}$ belongs to $M_F(\mathbb{R}^d)$, whereas $\mathscr{Z}_t^{(\eta_1,\eta_2)}$ belongs to $M_F(\mathbb{C}_d^{(0)})$. In addition, $Z_t^{(\eta_1,\eta_2)}$ is obtained from $\mathscr{Z}_t^{(\eta_1,\eta_2)}$ via projection, namely,

$$Z_t^{(\eta_1,\eta_2)} = \tilde{\pi}_t(\mathscr{Z}_t^{(\eta_1,\eta_2)}). \tag{1.6}$$

Finally, note that the operators of aggregation (cf. (1.5) and (1.5′)) and operators of projection $\tilde{\pi}$ and $\hat{\pi}$ (cf. (1.3), (1.3′), (1.4) and (1.6)) are in fact commutative.

For simplicity of notation, set $\eta_1 = \eta_2 = \eta$, and $X_t^{(\eta)} := X_t^{(\eta,\eta)}$. By analogy with the case of single-level branching, one can obtain that the empirical process $X_t^{(\eta)}$ possesses a limit (in both the weak and strong sense) as $\eta \to \infty$ that is a *two-level* $(2, d, \beta_1, \beta_2)$-*measure-valued process* X_t that starts with measure δ_{δ_0} at $t = 0$. Following Hochberg (1995), we will refer to such a limiting continuous-state process taking values in $M_F(M_F(\mathbb{R}^d))$ as a $(2, d, \beta_1, \beta_2)$-*super-2 process*. Note that one can also construct the $(2, d, \beta_1, \beta_2)$-*historical process* $\mathscr{H}_t$ by taking the limit as $\eta \to \infty$ of

$$\mathscr{H}_t^{(\eta)} := \mathscr{H}_t^{(\eta,\eta)}.$$

By analogy to [D, p. 8], we refer to the limits as $\eta \to \infty$ of these hierarchical branching-diffusing particle systems as *high-density* limits. This terminology is partially motivated by the fact that the limiting measure-valued processes are often so highly clumped that they live on sets of strictly smaller dimension (see Section 6 for more details).

Now, set

$$Z_t^{(\eta)} := Z_t^{(\eta,\eta)}, \tag{1.7}$$

and

$$\mathscr{Z}_t^{(\eta)} := \mathscr{Z}_t^{(\eta,\eta)}. \tag{1.7′}$$

Note that both $Z_t^{(\eta)}$ and $\mathscr{Z}_t^{(\eta)}$ play a very important role in the sequel.

2. Formulation of the main results

First, we present some notation. Let $\mathfrak{P}_\nu$ denote the probability law in $M_F(M_F(\mathbb{R}^d))$ of the super-2 process X_t at time instant t, where $\nu \in M_F(M_F(\mathbb{R}^d))$ denotes the initial measure. Let $\mathfrak{E}_\nu$ denote the conditional expectation with respect to $\mathfrak{P}_\nu$, given that the process starts with measure ν at $t = 0$.

Let $\mathscr{P}_\nu$ denote the probability law of X_t in $\mathbb{D}([0, \infty), M_F(M_F(\mathbb{R}^d)))$, and $\mathscr{Q}_{0,\nu}$ and $\mathscr{Q}_{0,\nu}^{(\eta)}$ denote the probability laws of the $(2, d, \beta_1, \beta_2)$-historical process $\mathscr{H}_t$ and of the corresponding BPS-2 historical process $\mathscr{H}_t^{(\eta)}$, respectively, where the subscript ν is to be understood as before.

We now define the *aggregated process* Z_t associated with the super-2 process X_t and taking values in $M_F(\mathbb{R}^d)$ (compare to formula (1.7) for the aggregated process

$Z_t^{(\eta)}$ associated with BPS-2) by

$$Z_t := \int_{M_F(\mathbb{R}^d)} \mu X_t(\mathrm{d}\mu).$$

The *historical aggregated process* $\mathscr{Z}_t$, which takes values in $M_F(\mathbb{C}_d^{(0)})$, is defined by

$$\mathscr{Z}_t := \int_{M_F(\mathbb{D}_d^{(1)})} \mu_t \mathscr{H}_t(\mathrm{d}\mu),$$

where $\mathscr{H}_t$ denotes the historical process associated with the $(2, d, \beta_1, \beta_2)$-super-2 process X_t that was introduced in Section 1. Note that the process $\mathscr{Z}_t(\mathbb{C}_d^{(0)})$ is the very important two-level analogue of the *total mass process* $M_t := U_t(\mathbb{R}^d)$ which is often used when studying super-1 processes (cf., e.g., [DV, Propositions 1.10 and 1.11]). The process $\mathscr{Z}_t(\mathbb{C}_d^{(0)})$ will emerge in some estimates of Section 5 of this work.

It is clear that the (non-historical) aggregated process Z_t and the historical aggregated process $\mathscr{Z}_t$ can also be obtained by taking high-density limits as $\eta \to \infty$ of $Z_t^{(\eta)}$ and of $\mathscr{Z}_t^{(\eta)}$, respectively (see (1.7) and (1.7′) for the definition of these processes).

We next define the real-valued process

$$r^{(1)}(t) := \inf\{R\colon S(U_t) \subseteq \overline{\mathbb{B}(0,R)}\},$$

where $S(Y)$ denotes the closed support of Y in $\mathbb{R}^d$, and $\overline{\mathbb{B}(0,R)}$ is the closed ball centered at the origin with radius R. This process was considered in [DV, p. 228] for describing the local propagation of mass from a point source. Analogously, we also introduce another real-valued process

$$r^{(2)}(t) := \inf\{R\colon S(Z_t) \subseteq \overline{\mathbb{B}(0,R)}\},$$

that describes the local propagation of mass from a point source for the aggregated process Z_t.

In the next theorem, we employ the Wiener probability measure $\mathbb{P}_0$ on the space of continuous trajectories that start from the origin, as well as the probability measures $\mathbf{P}_\mu$ and $\mathscr{P}_m$ introduced above.

Theorem 2.1. (i) *For any integer $d \geqslant 1$, there exist positive constants $C_1(d) \leqslant C_2(d)$ such that for any positive ε,*

$$\begin{aligned}\mathbb{P}_0\Bigg\{ & 1 + (C_1(d) - \varepsilon)\frac{\log\log 1/t}{\log 1/t} \leqslant \frac{\sup_{0 \leqslant s \leqslant 1-t} \sup_{0 < u \leqslant t} |w(s+u) - w(s)|}{\sqrt{2t \log 1/t}} \\ & \leqslant 1 + (C_2(d) + \varepsilon)\frac{\log\log 1/t}{\log 1/t}\ \textit{for all sufficiently small positive } t\Bigg\} \\ & = 1. \end{aligned} \tag{2.1}$$

(ii) *Let $m > 0$ be fixed. Then for any positive ε,*

$$\mathbf{P}_{m\cdot\delta_0}\left\{1 + \frac{1}{2}\left(\frac{d-2}{2(1/\beta)} - \varepsilon\right)\frac{\log\log 1/t}{\log 1/t} \leqslant \frac{\sup_{0\leqslant u\leqslant t} r^{(1)}(u)}{\sqrt{2(1/\beta)\,t\log(1/t)}}\right.$$

$$\left.\leqslant 1 + \frac{1}{2}\left(\frac{d/2+2}{1/\beta} + 2 + \varepsilon\right)\frac{\log\log 1/t}{\log 1/t} \text{ for all sufficiently small positive } t\right\}$$

$$= 1. \tag{2.1'}$$

(iii) *For any positive ε,*

$$\mathscr{P}_{\delta_{\delta_0}}\left\{1 + \frac{1}{2}\left(\frac{d-2}{2(1/\beta_1 + 1/\beta_2)} - \varepsilon\right)\frac{\log\log 1/t}{\log 1/t} \leqslant \frac{\sup_{0\leqslant u\leqslant t} r^{(2)}(u)}{\sqrt{2(1/\beta_1 + 1/\beta_2)\,t\log(1/t)}}\right.$$

$$\left.\leqslant 1 + \frac{1}{2}\left(\frac{d/2+2}{1/\beta_1 + 1/\beta_2} + 2 + \varepsilon\right)\frac{\log\log 1/t}{\log 1/t} \text{ for all sufficiently small positive } t\right\}$$

$$= 1. \tag{2.1''}$$

Remark 2.2. (i) Note that for the case $d = 1$, relationship (2.1) can be derived from Chung et al. (1959, Theorem 2) with $C_1(1) = C_2(1) = 5/4$. For $d \geqslant 2$, we can set $C_1(d) := d/4$, and $C_2(d) := 2 + d/4$. Recall that the proof of (2.1) can be found in [DHV3].

(ii) Note that this result contains a minor correction of the results stated in Theorems 1.3 and 1.7 of [DV], in which $\bar{\kappa}(\beta, d)$ should be $[(d/2) + 2]/(1/\beta) + 2$, and the coefficients of $[\log\log(1/t)]/\log(1/t)$ should be as stated above in formula (2.1′) of the present paper.

In particular, Theorem 2.1 yields the following corollary:

Corollary 2.3 (the exact almost-sure propagation of mass from a point source).

$$\text{(i)} \quad \mathbb{P}_0\left\{\lim_{t\downarrow 0} \frac{\sup_{0\leqslant s\leqslant 1-t}\sup_{0<u\leqslant t}|w(s+u) - w(s)|}{\sqrt{2t\cdot\log 1/t}} = 1\right\} = 1. \tag{2.2}$$

$$\text{(ii)} \quad \mathbf{P}_{m\cdot\delta_0}\left\{\lim_{t\downarrow 0} \frac{\sup_{0\leqslant u\leqslant t} r^{(1)}(u)}{\sqrt{2(1/\beta)\,t\log\dfrac{1}{t}}} = 1\right\} = 1. \tag{2.2'}$$

$$\text{(iii)} \quad \mathscr{P}_{\delta_{\delta_0}}\left\{\lim_{t\downarrow 0} \frac{\sup_{0\leqslant u\leqslant t} r^{(2)}(u)}{\sqrt{2(1/\beta_1 + 1/\beta_2)\,t\log(1/t)}} = 1\right\} = 1. \tag{2.2''}$$

Remark 2.4. Note that (2.2) is just the classical Lévy's result, whereas (2.2′) and (2.2″) can be viewed as *local-modulus-of-continuity-type* results (describing the exact almost-sure propagation of mass from a point source) for super-1 and super-2 processes, respectively. However, the local modulus of continuity for the Wiener process $w(\cdot)$ contains the iterated logarithm and hence has a different form than (2.2′) and (2.2″) (cf., e.g., Csörgő and Révész (1981, p. 41) for a heuristic explanation of this phenomenon). Namely, for any *fixed* $t_0 \geqslant 0$,

$$\mathbb{P}_0 \left\{ \lim_{t \downarrow 0} \frac{\sup_{0 \leqslant h \leqslant 1} |w(t_0 + h) - w(t_0)|}{\sqrt{2t \log\log(1/t)}} = 1 \right\} = 1.$$

The fact that the forms of the local moduli of continuity for super-1 and super-2 processes resemble the *global* (not local) modulus of continuity for the Wiener process is a consequence of the fact that even though we are considering the propagation of the closed supports of the super-1 and super-2 processes on *short time intervals*, here we must take into account the possibility of large increments of *many* individual Wiener processes, because of the branching. Thus, our situation is more similar in character to that of the global modulus of continuity for the Wiener process, where the possibility of large increments on a *large number* of short time intervals is considered, rather than to that of the local modulus of continuity, where the consideration centers about the possibility of large increments on a *single* short time interval. In this respect, we emphasize here that our results on the almost-sure rate of convergence in the *local* moduli for super-1 and super-2 processes are analogous to the result on the almost-sure rate of convergence in the *global* modulus for d-dimensional Wiener process (cf., e.g., Chung et al. (1959, Theorem 2)).

It should be mentioned that previous studies of the path properties of the processes $w(t)$, U_t and X_t have involved the use of some ideas first developed by Perkins in the setting of nonstandard analysis and later developed in the historical setting by Mueller and Perkins (1992). In particular, this approach involves the concept of so-called *bad realizations*, defined for $i \in \{1, 2\}$ and for integers $j > k$ by

$$\mathscr{G}^{(i)}_{n,j,k} := \{ y \in \mathbb{C}^{(0)}_d : |y(j \cdot 2^{-n}) - y(k \cdot 2^{-n})| > g^{(i)}_\kappa((j-k) \cdot 2^{-n}) \},$$

where the functions $g^{(i)}_\kappa(s)$ are defined as follows:

$$g^{(i)}_\kappa(0) := 0,$$

and for $t \in (0, \mathrm{e}^{-1}]$,

$$g^{(1)}_\kappa(t) := \sqrt{2(1 + 1/\beta) t (\log(1/t) + \kappa \log\log(1/t))}, \tag{2.3}$$

$$g^{(2)}_\kappa(t) := \sqrt{2(1 + 1/\beta_1 + 1/\beta_2) t (\log(1/t) + \kappa \log\log(1/t))}. \tag{2.3′}$$

Now, we introduce a new concept of the *historical aggregated mass* of "bad" realizations of historical BPS-2 and $(2, d, \beta_1, \beta_2)$-historical processes in order to study the path properties of two-level superprocesses. These historical aggregated masses are respectively defined by

$$\mathscr{K}^{(n)}_{n,j,k}(t) := \mathscr{Z}^{(n)}_{j \cdot 2^{-n} + t}(\mathscr{G}^{(2)}_{n,j,k}) \quad \text{and} \quad \mathscr{K}_{n,j,k}(t) := \mathscr{Z}_{j \cdot 2^{-n} + t}(\mathscr{G}^{(2)}_{n,j,k}).$$

Note that $\mathscr{K}_{n,j,k}(t)$ generalizes the real-valued process

$$\mathscr{M}_{n,j,k}(t) := H_{j\cdot 2^{-n}+t}(\mathscr{G}^{(1)}_{n,j,k}),$$

introduced in Mueller and Perkins (1992, pp. 340–341), that gives the values taken by the $(2, d, \beta)$-historical process H_t on the sets $\mathscr{G}^{(1)}_{n,j,k}$.

In addition, we need to modify these processes slightly in order to formulate the *local* modulus-of-continuity-type results for super-1 and super-2 processes (see formula (2.8) and Theorem 2.7 below). To this end, for $i \in \{1,2\}$ set

$$h^{(i)}_{\kappa}(0) := 0,$$

and for $t \in (0, \mathrm{e}^{-1}]$, set

$$h^{(1)}_{\kappa}(t) := \sqrt{2(1/\beta)t(\log(1/t) + \kappa \log\log(1/t))}. \tag{2.4}$$

$$h^{(2)}_{\kappa}(t) := \sqrt{2(1/\beta_1 + 1/\beta_2)t(\log(1/t) + \kappa \log\log(1/t))}. \tag{2.4$'$}$$

For $i \in \{1,2\}$, let $\hat{\mathscr{G}}^{(i)}_{n,j,k}$ and $\hat{\mathscr{M}}^{(i)}_{n,j,k}$ be defined as $\mathscr{G}^{(i)}_{n,j,k}$ and $\mathscr{M}^{(i)}_{n,j,k}(t)$, with the function $g^{(i)}_{\kappa}(\cdot)$ replaced by the function $h^{(i)}_{\kappa}(\cdot)$. Also, set

$$K^{(i)}_{\kappa}(\varepsilon, u) := \{ y \in \mathbb{C}^0_d : |y(u) - y(s)| \leqslant g^{(i)}_{\kappa}(u-s) \text{ for all } 0 \leqslant s < u;\ u - s \leqslant \varepsilon \}, \tag{2.5}$$

where the functions $g^{(i)}_{\kappa}(\cdot)$ are defined in terms of (2.3) and (2.3′). Note that the sets $K^{(i)}_{\kappa}(\varepsilon, u)$ of paths are used in Theorem 2.5 for the description of the almost-sure propagation of the closed supports of the corresponding historical processes.

Now, given a $(2, d, \beta)$-super-1 process U_t and its sample path ω that belongs to $\mathbb{D}([0, \infty), M_{\mathrm{F}}(\mathbb{R}^d))$, define

$$T_1(\omega) := \inf\{t \colon \langle U_t(\omega), \mathbb{R}^d \rangle = 0\}.$$

In addition, given a $(2, d, \beta_1, \beta_2)$-super-2 process X_t and its sample path ω that belongs to $\mathbb{D}([0, \infty), M_{\mathrm{F}}(M_{\mathrm{F}}(\mathbb{R}^d)))$, define

$$T_2(\omega) := \inf\{t \colon \langle Z_t(\omega), \mathbb{R}^d \rangle = 0\}.$$

Note that both the super-1 process U_t and the super-2 process X_t become extinct in finite time. For U_t, this follows, e.g., from the fact that the probability of non-extinction of the total mass process $M_t := U_t(\mathbb{R}^d)$ (where $\mathbb{R}^d$ is the state space) decays like $t^{-1/\beta}$ as $t \to \infty$ (cf., e.g., [DV, Proposition 1.10]) and a subsequent application of the Borel–Cantelli lemma. In addition, since the behavior of the total mass process M_t does not depend at all on the state space, we conclude that the extinction property of the super-2 process X_t follows from the extinction of the corresponding super-1 process $\tilde{U}_t$ with state space $M_{\mathrm{F}}(\mathbb{R}^d)$. Hence, the stopping times T_1 and T_2 are finite with probability 1.

The following result provides the global moduli of continuity for the closed supports $S(H_t)$ and $S(\mathscr{Z}_t)$ of the $(2, d, \beta)$-historical process H_t and of the historical aggregated process $\mathscr{Z}_t$ in $\mathbb{R}^d$, respectively.

Theorem 2.5. (i) *Let* $\mu \in M_F(\mathbb{R}^d)$ *and*

$$\kappa > \kappa(\beta, d) = \frac{d + 8 + 4/\beta}{2(1 + 1/\beta)}.$$

Then for $\mathbb{Q}_{0,\mu}$*-a.s.* ω, *there exists a* $\delta_1(\omega, \kappa) > 0$ *such that for any real* $0 < t \leqslant T_1(\omega)$,

$$S(H_t(\omega)) \subseteq K_\kappa^{(1)}(\delta_1(\omega,\kappa), t) \subseteq K_\kappa^{(1)}(\delta_1(\omega,\kappa), T_1(\omega)). \tag{2.6}$$

(ii) *Let* $m \in M_F(M_F(\mathbb{R}^d))$ *and*

$$\kappa > \kappa(\beta_1, \beta_2, d) = \frac{d + 8 + 4/\beta_1 + 4/\beta_2}{2(1 + 1/\beta_1 + 1/\beta_2)}.$$

Then for $\mathscr{Q}_{0,m}$*-a.s.* ω, *there exists a* $\delta_2(\omega, \kappa) > 0$ *such that for any real* $0 < t \leqslant T_2(\omega)$,

$$S(\mathscr{X}_t(\omega)) \subseteq K_\kappa^{(2)}(\delta_2(\omega,\kappa), t) \subseteq K_\kappa^{(2)}(\delta_2(\omega,\kappa), T_2(\omega)). \tag{2.6'}$$

In particular, Theorem 2.5 yields the following corollary related to the global moduli of continuity for super-1 and super-2 processes:

Corollary 2.6. (i) *Let* $\mu \in M_F(\mathbb{R}^d)$ *and*

$$\kappa > \kappa(\beta, d) = \frac{d + 8 + 4/\beta}{2(1 + 1/\beta)}.$$

Let $T > 0$ *be fixed. Then for* $\mathbf{P}_\mu$*-a.e.* ω, *there exists a* $\bar{\delta}_1(\omega, \kappa) > 0$ *such that if* $0 \leqslant s, t \leqslant T$ *satisfy* $0 < t - s < \bar{\delta}_1$, *then*

$$S(U_t) \subseteq S(U_s)^{g_\kappa^{(1)}(t-s)}. \tag{2.7}$$

(ii) *Let* $m \in M_F(M_F(\mathbb{R}^d))$ *and*

$$\kappa > \kappa(\beta_1, \beta_2, d) = \frac{d + 8 + 4/\beta_1 + 4/\beta_2}{2(1 + 1/\beta_1 + 1/\beta_2)}.$$

Let $T > 0$ *be fixed. Then for* $\mathscr{P}_m$*-a.e.* ω, *there exists a* $\bar{\delta}_2(\omega, \kappa) > 0$ *such that if* $0 < t - s < \bar{\delta}_2$, *then*

$$S(Z_t) \subseteq S(Z_s)^{g_\kappa^{(2)}(t-s)}. \tag{2.7'}$$

Here

$$A^\varepsilon := \{x \in \mathbb{R}^d\colon \operatorname{dist}(x, A) \leqslant \varepsilon\}$$

denotes the ε*-neighborhood of the set* A.

We now present the local analogue of Theorem 2.5 that will be used in the proof of parts (ii) and (iii) of Theorem 2.1. To this end, we slightly modify our notation. For $i \in \{1, 2\}$ and for $s \geqslant 0$, let

$$\hat{K}_\kappa^{(i)}(\varepsilon, s) := \{y \in \mathbb{C}_d^0\colon\ |y(u) - y(s)| \leqslant h_\kappa^{(i)}(u - s) \text{ for all } s \leqslant u \leqslant s + \varepsilon\}, \tag{2.8}$$

where the functions $h_\kappa^{(i)}(\cdot)$ are defined in terms of (2.4) and (2.4').

Theorem 2.7. (i) *Let* $\mu \in M_F(\mathbb{R}^d)$ *and*

$$\kappa > \hat{\kappa}(\beta, d) := \frac{d/2 + 2}{1/\beta} + 2.$$

Then for each fixed $t \geqslant 0$ *and for* $\mathbb{Q}_{0,\mu}$*-a.e.* ω, *there exists a* $\delta_*^{(1)}(\omega, \kappa) > 0$ *such that for any* $t \leqslant s \leqslant t + \delta_*^{(1)}$,

$$S(H_s(\omega)) \subseteq \hat{K}_\kappa(\delta_*^{(1)}(\omega, \kappa), t). \tag{2.9}$$

(ii) *Let* $m \in M_F(M_F(\mathbb{R}^d))$ *and*

$$\kappa > \hat{\kappa}(\beta_1, \beta_2, d) := \frac{d/2 + 2}{1/\beta_1 + 1/\beta_2} + 2.$$

Then for each fixed $t \geqslant 0$ *and for* $\mathbb{Q}_{0,m}$*-a.e.* ω, *there exists a* $\delta_*^{(2)}(\omega, \kappa) > 0$ *such that for any* $t \leqslant s \leqslant t + \delta_*^{(2)}$,

$$S(\mathscr{X}_s(\omega)) \subseteq \hat{K}_\kappa(\delta_*^{(2)}(\omega, \kappa), t). \tag{2.9'}$$

In particular, Theorem 2.7 yields

Corollary 2.8. (i) *Let* $\mu \in M_F(\mathbb{R}^d)$ *and* $\kappa > (d/2 + 2)/(1/\beta) + 2$. *Then for each fixed* $t \geqslant 0$ *and for* $\mathbf{P}_\mu$*-a.e.* ω, *there exists a* $\bar{\delta}_*^{(1)}(\omega, \kappa)$ *such that if* $0 < s < \bar{\delta}_*^{(1)}$, *then*

$$S(U_{t+s}) \subseteq S(U_t)^{h_\kappa^{(1)}(s)}. \tag{2.10}$$

(ii) *Let* $m \in M_F(M_F(\mathbb{R}^d))$ *and*

$$\kappa > \frac{d/2 + 2}{1/\beta_1 + 1/\beta_2} + 2.$$

Then for each fixed $t \geqslant 0$ *and for* $\mathscr{P}_m$*-a.e.* ω, *there exists a* $\bar{\delta}_*^{(2)}(\omega, \kappa)$ *such that if* $0 < s < \bar{\delta}_*^{(2)}$, *then*

$$S(Z_{t+s}) \subseteq S(Z_t)^{h_\kappa^{(2)}(s)}. \tag{2.10'}$$

3. Analytical representation of superprocesses

In this section, we briefly describe an analytical approach to the study of super-1 and super-2 processes. Let $\{T_t: t \geqslant 0\}$, $T_t: \mathbb{C}(M_F(\mathbb{R}^d)) \to \mathbb{C}(M_F(\mathbb{R}^d))$ be a semigroup, associated with the semigroup $\{V_t: t \geqslant 0\}$ of the single-level $M_F(\mathbb{R}^d)$-valued branching-diffusion process via the relationship

$$T_t \exp\{-\langle \phi, \cdot \rangle\} := \exp\{-\langle V_t \phi, \cdot \rangle\}, \tag{3.1}$$

where

$$V_t\phi(x) := S_t\phi(x) - \frac{1}{1+\beta}\int_0^t S_u[(V_{t-u}\phi)^{1+\beta}]\,\mathrm{d}u. \tag{3.2}$$

Here and below, $\langle\phi,\mu\rangle := \int \phi\,\mathrm{d}u$, the semigroup $\{S_t;\ t \geqslant 0\}$ is associated with the infinitesimal generator $\Delta/2$ of the Wiener process, and $\beta \in (0,1]$. For a comprehensive review of the characterization of single-level super-1 processes, see [D].

Recall (see the beginning of Section 2) that $\mathfrak{P}_\nu$ denotes the probability law in $M_{\mathrm{F}}(M_{\mathrm{F}}(\mathbb{R}^d))$ of the super-2 process X_t at time instant t, where $\nu \in M_{\mathrm{F}}(M_{\mathrm{F}}(\mathbb{R}^d))$ denotes the initial measure, and that $\mathfrak{E}_\nu$ denotes the conditional expectation with respect to $\mathfrak{P}_\nu$, given that the process starts with measure ν at $t = 0$.

Note that a super-2 process X_t can be viewed as a super-superprocess, that is, a superprocess arising as the limit of a certain branching-diffusing particle system in which the "particles" live in the state space $M_{\mathrm{F}}(\mathbb{R}^d)$, while their motion mechanism is determined by a super-1 process. (This point of view is pursued in Section 3.4.2 of Bojdecki and Gorostiza (1995).) Then, by analogy with the $(2,d,\beta)$-super-1 process U_t, the $(2,d,\beta_1,\beta_2)$-super-2 process X_t can be characterized via the Laplace functional $L_{t,\nu}(J)$ given by

$$\begin{aligned} L_{t,\nu}(J) &:= \mathfrak{E}_\nu\left\{\exp\left(-\int_{M_{\mathrm{F}}(\mathbb{R}^d)} J(\mu)X(t,\mathrm{d}\mu)\right)\right\} \\ &= \exp\left(-\int u(t,\mu)\,\nu(\mathrm{d}\mu)\right). \end{aligned}$$

where $u(t,\mu)$ satisfies the integral equation

$$u(t,\mu) = T_t u(0,\mu) - \frac{1}{1+\beta_2}\int_0^t [T_{t-s}u^{1+\beta_2}(s,\cdot)](\mu)\,\mathrm{d}s, \qquad u(0,\mu) = J(\mu)$$

for $J(\mu) = f(\langle\phi,\mu\rangle)$, where ϕ is $\mathbb{R}^d$-valued continuous and f is real-valued bounded continuous. Here, the semigroup T_t is expressed in terms of (3.1) and (3.2) with $\beta = \beta_1$.

Note that interplays between certain properties of super-1 processes and those of solutions of certain nonlinear differential equations have been studied in a number of works (cf., e.g., Dynkin, 1993). Some relations between path properties of super-2 processes and properties of a certain infinite-dimensional partial differential equation have been established in [DHV2, Section 3]. In addition, the analytical method sketched in this section has been successfully applied for the derivation of many interesting properties of super-2 processes, such as extinction, Hausdorff dimension, self-similarity and persistence.

4. Cluster structure of superprocesses

We first proceed with arguments that are similar to those used in the proofs of Theorem 2.1 of Tribe (1989) and of [DV, Theorem 3.1] in the single-level setting. Note that the initial distribution m_η of the empirical process $X_t^{(\eta)}$ (described in Section 1) can be represented as

$$m_\eta := X_0^{(\eta)} \stackrel{d}{=} \sum_{i=1}^{\Pi(\eta)} \frac{1}{\eta} \delta_{\sum_{j=1}^{\Pi_i(\eta)} \frac{1}{\eta} \delta_j}, \tag{4.1}$$

where all the δ's denotes δ-functions located at the origin. In particular, (4.1) implies that the sequence $\{m_\eta\}$ of random measures converges to δ_{δ_0} as $\eta \to \infty$ in a strong sense.

Recall that one of the key tools in the theory of branching-diffusion processes is the consideration of cluster structures. In the case of single-level branching-diffusing particle systems, we refer to the collection of all the descendants of a single particle from the initial set that are alive at time t as the *cluster of age* t generated by that initial, or *tagged*, particle. It is natural that in the case of *hierarchical* branching-diffusing particles, one has to take into account a *hierarchical* cluster structure. We now proceed with its description.

Let $K_{\eta,\beta_2}^{(2)}(t)$ denote the number of initial superparticles of $X_t^{(\eta)}$ which have living descendants at time t. Clearly,

$$K_{\eta,\beta_2}^{(2)}(0) \stackrel{d}{=} \Pi(\eta). \tag{4.2}$$

It is natural to refer to $K_{\eta,\beta_2}^{(2)}(t)$ as the *number of level-2 clusters of age t*; each individual level-2 cluster is generated by a surviving superparticle from the initial set. In particular, the random measure of the empirical process $X_t^{(\eta)}$ at time t can be represented as the union of the random number $K_{\eta,\beta_2}^{(2)}(t)$ of level-2 clusters of age t. Note that some of the superparticles might contain no level-1 particles; hereafter, we refer to such superparticles as *null superparticles*. Obviously, $K_{\eta,\beta_2}^{(2)}(t)$ is in fact the random sum of a Poisson number $\Pi(\eta)$ of i.i.d. 0/1-valued Bernoulli random variables, and the probability of success $Q_{\eta,\beta_2}^{(2)}(t)$ in a single trial (i.e., survival of descendants of an individual superparticle from the initial set at instant t) satisfies the relationship

$$Q_{\eta,\beta_2}^{(2)}(t) = \left(1 + \frac{\beta_2}{\beta_2 + 1} \cdot t \cdot \eta^{\beta_2}\right)^{-1/\beta_2} \tag{4.3}$$

(cf., e.g., Zolotarev, 1957, Section 5) or [DV, formula (1.12)]). In addition, $K_{\eta,\beta_2}^{(2)}(t)$ has a Poisson distribution with parameter $\eta \cdot Q_{\eta,\beta_2}^{(2)}(t)$. Indeed, it can be shown that for any $t > 0$ and for any integer $l \geqslant 0$,

$$\mathfrak{P}_{m_\eta}\{K_{\eta,\beta_2}^{(2)}(t) = l\} = \sum_{k=l}^{\infty} \mathrm{e}^{-\eta} \frac{(\eta)^k}{k!} \binom{k}{l} (Q_{\eta,\beta_2}^{(2)}(t))^l (1 - Q_{\eta,\beta_2}^{(2)}(t))^{k-l}. \tag{4.4}$$

Now, we derive a useful representation for the Laplace transform $\psi_{\eta,t}^{(\beta_2)}(\cdot)$ of $K_{\eta,\beta_2}^{(2)}(t)$. Applying (4.4), we get that for any real $s \geqslant 0$,

$$\begin{aligned} \psi_{\eta,t}^{(\beta_2)}(s) &= \sum_{l=0}^{\infty} \mathrm{e}^{-ls} \cdot \left(\sum_{k=l}^{\infty} \mathrm{e}^{-\eta} \frac{(\eta)^k}{k!} \binom{k}{l} (Q_{\eta,\beta_2}^{(2)}(t))^l (1 - Q_{\eta,\beta_2}^{(2)}(t))^{k-l} \right) \\ &= \mathrm{e}^{-\eta} \sum_{k=0}^{\infty} \frac{(\eta)^k}{k!} \left(\sum_{l=0}^{k} \mathrm{e}^{-ls} \binom{k}{l} (Q_{\eta,\beta_2}^{(2)}(t))^l (1 - Q_{\eta,\beta_2}^{(2)}(t))^{k-l} \right). \end{aligned} \tag{4.5}$$

Note that the sum within the braces on the right-hand side of (4.5) is in fact the Laplace transform of the sum of the number k of i.i.d. 0/1-valued Bernoulli random variables with probability of success in a single trial equal to $Q^{(2)}_{\eta,\beta_2}(t)$. Therefore,

$$\sum_{l=0}^{k} \mathrm{e}^{-ls}\binom{k}{l}(Q^{(2)}_{\eta,\beta_2}(t))^l(1 - Q^{(2)}_{\eta,\beta_2}(t))^{k-l} = (1 - Q^{(2)}_{\eta,\beta_2}(t)(1 - \mathrm{e}^{-s}))^k.$$

Combining the above representation with (4.5), we obtain the result that for any real $s \geqslant 0$,

$$\begin{aligned}\psi^{(\beta_2)}_{\eta,t}(s) &= \mathrm{e}^{-\eta}\cdot\sum_{k=0}^{\infty}\frac{1}{k!}\cdot\{\eta\cdot(1 - Q^{(2)}_{\eta,\beta_2}(t)\cdot(1 - \mathrm{e}^{-s}))\}^k \\ &= \exp\{\eta\cdot Q^{(2)}_{\eta,\beta_2}(t)\cdot(\mathrm{e}^{-s} - 1)\}. \qquad (4.6)\end{aligned}$$

In turn, (4.6) implies that $K^{(2)}_{\eta,\beta_2}(t)$ has a Poisson distribution with parameter $\eta\cdot Q^{(2)}_{\eta,\beta_2}(t)$.

Now, fix the ith superparticle alive at time t (which might be a null superparticle). Clearly, it is comprised of a random number $K^{(1)}_{\eta,\beta_1}(t;i)$ of level-1 clusters. Each level-1 cluster is comprised of the surviving descendants of an individual particle from the initial set. By analogy with (4.1) and (4.2), we get that

$$K^{(1)}_{\eta,\beta_1}(0;i) \stackrel{\mathrm{d}}{=} \Pi_i(\eta). \qquad (4.7)$$

Since each newly born superparticle is a copy of its parent superparticle, by analogy with the arguments leading to (4.3), it follows from (4.7) that $K^{(1)}_{\eta,\beta_1}(t;i)$ can be viewed as the random sum of Poisson numbers $\Pi_i(\eta)$ of i.i.d. 0/1-valued Bernoulli random variables, where the probability of success $Q^{(1)}_{\eta,\beta_1}(t)$ in a single trial (i.e., survival of descendants of an individual particle belonging to the ith superparticle at instant t) satisfies relationship (4.3) with the change of 2's for 1's. In other words, in order to determine the distribution of the number of level-1 clusters within an *individual superparticle* at time t, we can suppress the second-level branching and regard this superparticle as a single-level branching particle system.

By analogy with (4.6), one can get that the distribution of $K^{(1)}_{\eta,\beta_1}(t;i)$ is Poisson with parameter $\eta\cdot Q^{(1)}_{\eta,\beta_1}(t)$. Indeed, this follows from the fact that the Laplace transform $\psi^{(\beta_1)}_{\eta,t}(\cdot)$ of $K^{(1)}_{\eta,\beta_1}(t;i)$ satisfies the relationship

$$\begin{aligned}\psi^{(\beta_1)}_{\eta,t}(s) &= \mathrm{e}^{-\eta}\cdot\sum_{k=0}^{\infty}\frac{1}{k!}[\eta(1 - Q^{(1)}_{\eta,\beta_1}(t)(1 - \mathrm{e}^{-s}))]^k \\ &= \exp\{\eta\cdot Q^{(1)}_{\eta,\beta_1}(t)(\mathrm{e}^{-s} - 1)\}. \qquad (4.8)\end{aligned}$$

Recall that by our choice of the initial distribution, any initial superparticle consists of $\Pi_i(\eta)$ particles, where $1 \leqslant i \leqslant \Pi(\eta)$. On the other hand, superparticles alive at time t consist of random numbers $K^{(1)}_{\eta,\beta_1}(t;i)$ of level-1 clusters, and the random numbers $K^{(1)}_{\eta,\beta_1}(t;1)$, $K^{(1)}_{\eta,\beta_1}(t;2)$, ... are identically distributed (though some of them – those related to superparticles that belong to the same level-2 clusters – are dependent).

However, random numbers $K^{(1)}_{\eta,\beta_1}(t;i_1)$ and $K^{(1)}_{\eta,\beta_1}(t;i_2)$ of level-1 clusters related to superparticles belonging to *different level*-2 *clusters* are independent. The main idea behind the derivation of the lower bound for the empirical process $X^{(\eta)}_t$ consists in choosing one and only one representative (hereafter referred to as the *tagged superparticle*) from each surviving level-2 cluster, and then one and only one representative (hereafter referred to as a *tagged particle*) from each surviving level-1 cluster (which belongs to a tagged superparticle). Note that any tagged particle belongs to a certain tagged superparticle. It is clear that the exclusion from our consideration of non-tagged superparticles and non-tagged particles (within tagged superparticles) will give us a conservative lower bound. On the other hand, this simplification will enable us to exploit the independence of tagged superparticles (belonging to *different* level-2 clusters of age t) as well as the independence of tagged particles within a tagged superparticle (belonging to *different* level-1 clusters of age t). In addition, the independence of tagged superparticles and the independence of tagged particles within a superparticle, along with the above representations in terms of random sums of i.i.d. 0/1-valued Bernoulli random variables, enable us to derive the following useful representation for the number of *tagged level*-1 *clusters of age t* of the empirical process $X^{(\eta)}_t$ (cf. Proposition 4.1). Recall that a similar idea was used by Tribe (1989, Theorem 2.1) and by [DV, Theorem 3.1] in the single-level setting.

Proposition 4.1. *Consider the empirical process $X^{(\eta)}_t$ described in Section* 1 *with branching mechanisms on the first and second hierarchical levels governed by* (1.2) *with $\beta = \beta_1$ and $\beta = \beta_2$, respectively. Then for any fixed real $t > 0$, the random sum*

$$\mathscr{L}_{\eta,\beta_1,\beta_2}(t) := \sum_{i=1}^{K^{(2)}_{\eta,\beta_2}(t)} K^{(1)}_{\eta,\beta_1}(t;i)\,, \tag{4.9}$$

which represents the number of tagged level-1 clusters of age t *of the empirical process $X^{(\eta)}_t$, has the following Laplace transform:*

$$\psi^{(\beta_1,\beta_2)}_{\eta,t}(s) = \psi^{(\beta_2)}_{\eta,t}\left(\log\frac{1}{\psi^{(\beta_1)}_{\eta,t}(s)}\right).$$

Proof. Straightforward. □

It turns out that a combination of Proposition 4.1 with (4.2), (4.6), (4.9) and the characterization of weak convergence of non-negative random variables in terms of pointwise convergence of their Laplace transforms yield the following result:

Corollary 4.2. *For any fixed real $t > 0$, the random sum $\mathscr{L}_{\eta,\beta_1,\beta_2}(t)$ converges weakly as $\eta \to \infty$ to a random variable $\mathscr{L}_{\infty,\beta_1,\beta_2}(t)$ having the Laplace transform*

$$f^{(\beta_1,\beta_2)}_t(s) := \mathbb{E}\{\exp\{-s\cdot\mathscr{L}_{\infty,\beta_1,\beta_2}(t)\}\} = f^{(\beta_2)}_t\left(\log\frac{1}{f^{(\beta_1)}_t(s)}\right),$$

where

$$f_t^{(\beta_i)}(s) := \exp\{(\beta_i \cdot t/(\beta_i + 1))^{-1/\beta_i}(\mathrm{e}^{-s} - 1)\}.$$

Remark 4.3. Note that the random variable $\mathscr{L}_{\infty,\beta_1,\beta_2}(t)$ can be viewed as the number of *tagged level*-1 *clusters of age* t of the super-2 process X_t, which in our case is compound Poisson. In this respect, note that Corollary 4.2 is in the same spirit as Corollary 11.5.3b of [D], although the latter is related to the case of single-level branching with a *Poisson* rather than a *compound Poisson* cluster representation.

Now, let us present the almost-sure lower estimate for the local propagation of mass from a point source for super-2 processes. To this end, we introduce the following auxiliary family of continuous increasing functions on the interval $[0, \mathrm{e}^{-1}]$:

$$v_{\beta_1,\beta_2,\varepsilon}(s) := \sqrt{2\left(\frac{1}{\beta_1} + \frac{1}{\beta_2}\right)s\left(\log\frac{1}{s} + \left(\frac{d-2}{2(1/\beta_1 + 1/\beta_2)} - \varepsilon\right)\log\log\frac{1}{s}\right)}$$

for $s \in (0, \mathrm{e}^{-1}]$,

and

$$v_{\beta_1,\beta_2,\varepsilon}(0) := 0$$

(compare to the definition of the function $h_\kappa^{(2)}(\cdot)$ in formula (2.4′)). In what follows, the parameter ε is assumed to take any fixed, sufficiently small positive value less than one if $d \leqslant 2$, and less than $(d-2)/[2(1/\beta_1 + 1/\beta_2)]$ if $d \geqslant 3$.

Theorem 4.4. *For any* $\varepsilon > 0$,

$$\mathscr{P}_{\delta_{\delta_0}}\left\{\sup_{0 \leqslant u \leqslant t} r^{(2)}(u) \geqslant v_{\beta_1,\beta_2,\varepsilon}(t) \textit{ for all sufficiently small positive } t\right\} = 1. \tag{4.10}$$

Proof (*sketch*). Let us describe our method on a heuristic level. It can be shown that almost-sure lower estimates for the propagation of support can be derived from lower estimates for

$$\mathscr{P}_{\delta_{\delta_0}}\left\{\int \mu(\overline{\mathbb{B}(0,R)}^c)\, X_s(\mathrm{d}\mu) > 0 \text{ for some } 0 \leqslant s \leqslant t\right\}$$

by the use of Borel–Cantelli arguments. In turn, estimating the above probability is equivalent to the derivation of upper estimates for the probability of the event that the support of the $(2, d, \beta_1, \beta_2)$-super-2 process X_t will remain inside the closed ball $\overline{\mathbb{B}(0,R)}$ during the time period $[0, t]$. This approach can be developed by the use of BPS-2 approximation, Proposition 4.1 and Borel–Cantelli arguments. The main idea lies in the simultaneous consideration of the $(2, d, \beta_1, \beta_2)$-super-2 process X_t and the empirical process $X_t^{(\eta)}$. In particular, Proposition 4.1 enables one to apply the Borel–Cantelli lemma in order to establish convergence of a certain series related to the $(2, d, \beta_1, \beta_2)$-super-2 process X_t. Indeed, such probabilities can be obtained as limits as $\eta \to \infty$ of probabilities of certain events related to the corresponding empirical process $X_t^{(\eta)}$. Subsequently, the latter probabilities can be estimated from above in terms of $\psi_{\eta,t}^{(\beta_1,\beta_2)}(\cdot)$, and this function can also be properly estimated. In other words, we can establish almost-sure lower bounds (i.e., find proper lower functions) for the asymptotic behavior of $\sup_{0 \leqslant u \leqslant t} r^{(2)}(u)$ as $t \to 0$. In addition, only purely probabilistic arguments and some well-known properties of d-dimensional Wiener process need be used.

The proof itself develops along the same lines as that of [DV, Theorem 3.1] and actually repeats that of [DHV1, Theorem 2.1] and therefore is omitted. □

5. Almost-sure upper estimates

In this section, we first prove Theorems 2.5(ii) and 2.7(ii), which provide global and local moduli of continuity for the $(2, d, \beta_1, \beta_2)$-*historical* process $\mathscr{H}_t$ associated with the $(2, d, \beta_1, \beta_2)$-super-2 process X_t. Then, at the end of this section, we combine the results of Theorems 2.7(ii) and 4.4 to derive Theorem 2.1(iii). Note that in order to prove Theorems 2.5(ii) and 2.7(ii), we need to construct a sequence of historical processes $\mathscr{H}_t^{(\eta)}$ associated with the BPS-2 $X_t^{(\eta)}$ described in Section 1. Hereafter, we refer to $\mathscr{H}_t^{(\eta)}$ as the *BPS*-2 *historical processes.*

Note that the historical processes $\mathscr{H}_t$ and $\mathscr{H}_t^{(\eta)}$ take values in the space $M_{\mathrm{F}}(\mathbb{D}_d^{(1)})$, and that the historical aggregated processes $\mathscr{Z}_t^{(\eta)}$ and $\mathscr{Z}_t$ take values in the space $M_{\mathrm{F}}(\mathbb{C}_d^{(0)})$. Recall that the sequence $\mathscr{H}_t^{(\eta)}$ converges weakly to $\mathscr{H}_t$ as $\eta \to \infty$ (c.f., e.g., Wu (1992, Chapter 4) and Wu (1993) for the special case $\beta_1 = \beta_2 = 1$ in a non-historical setting; the general case $\beta_1 \in (0,1]$ and $\beta_2 \in (0,1]$ is obtained by slightly modifying arguments of the special case). In particular, this approach enables one to derive moduli of continuity by taking limits as $\eta \to \infty$ of upper estimates for probability distributions of the BPS-2 historical processes $\mathscr{H}_{\cdot}^{(\eta)}$.

Now, we briefly describe the main methods used in this section. Our first auxiliary result, Lemma 5.1, in a modification of [DV, Lemma 2.1]. Also, note that Lemma 5.2 of this section provides an upper estimate for probabilities of "bad" realizations of BPS-2 historical processes $\mathscr{H}_t^{(\eta)}$, in terms of the product of the probability of survival of particles from the initial set and the measure of "bad" realizations of the d-dimensional Wiener process. This lemma is in the same spirit as estimate (3.16) of Mueller and Perkins (1992). Subsequently, an application of Lemma 5.2, along with a slight modification of the arguments used in [DV, Section 2] for the derivation of a refinement of the global modulus of continuity for super-1 processes, enables us to establish the results of Theorems 2.5(ii) and 2.7(ii).

We now proceed with the formulation of a purely analytical lemma, which describes properties of the functions $g_\kappa^{(2)}(\cdot)$ and $h_\kappa^{(2)}(\cdot)$ and which is used for the proofs of Theorems 2.5(ii) and 2.7(ii). To this end, we introduce the function $\{n_0(u)$:

$0 < u < (\log 2)^2/2\}$ which is defined as the unique integer $n \geqslant 1$ such that

$$(\log 2)^2 (n+1)^2/2^{n+1} \leqslant u < (\log 2)^2 n^2/2^n.$$

Also, note that $n_0(u) \to \infty$ as $u \to 0$, and recall that

$$\kappa(\beta_1, \beta_2, d) = \frac{d + 8 + 4/\beta_1 + 4/\beta_2}{2(1 + 1/\beta_1 + 1/\beta_2)}$$

and

$$\hat{\kappa}(\beta_1, \beta_2, d) = \frac{d/2 + 2}{1/\beta_1 + 1/\beta_2} + 2$$

(see the formulation of Theorems 2.5(ii) and 2.7(ii)).

Lemma 5.1. *Let the function $g_\kappa^{(2)}(\cdot)$ be defined by (2.3′), the function $h_\kappa^{(2)}(\cdot)$ be defined by (2.4′),*

$$\kappa_1 := (\kappa + \kappa(\beta_1, \beta_2, d))/2,$$

and

$$\kappa_2 := (\kappa + \hat{\kappa}(\beta_1, \beta_2, d))/2.$$

(i) *If $\kappa > \kappa(\beta_1, \beta_2, d)$, then for any fixed positive constant C, there exists a positive integer $n(g_\kappa^{(2)}, C)$ such that*

$$g_\kappa^{(2)}(u) \geqslant g_{\kappa_1}^{(2)}(u) + C \cdot g_{\kappa_1}^{(2)}(1/2^{n_0(u)}) \tag{5.1}$$

for all u with $n_0(u) \geqslant n(g_\kappa^{(2)}, C)$.

(ii) *If $\kappa > \hat{\kappa}(\beta_1, \beta_2, d)$ then for any fixed positive constant C, there exists a positive integer $n(h_\kappa^{(2)}, C)$ such that*

$$h_\kappa^{(2)}(u) \geqslant h_{\kappa_2}^{(2)}(u) + C \cdot h_{\kappa_2}^{(2)}(1/2^{n_0(u)}) \tag{5.1′}$$

for all u with $n_0(u) \geqslant n(h_\kappa^{(2)}, C)$.

(iii) *For any $\kappa > \kappa(\beta_1, \beta_2, d)$, there exists a constant $C(\beta_1, \beta_2, d, \kappa)$ such that for all $n \geqslant 2$,*

$$\sum_{l=n+1}^{\infty} g_\kappa^{(2)}(1/2^l) \leqslant C(\beta_1, \beta_2, d, \kappa)\, g_\kappa^{(2)}(1/2^n). \tag{5.2}$$

Proof. Similar to that of [DV, Lemma 2.1] and therefore is omitted. □

Lemma 5.2. (Compare to formula (3.16) of Mueller and Perkins (1992)). *For any fixed real $t \geqslant 2^{-n}$,*

$$\begin{aligned}
&\mathscr{Q}_{0,m_\eta}^{(\eta)}(\mathscr{X}_{n,j,k}^{(\eta)}(t) > 0 \mid \mathscr{Z}_{j/2^n}^{(\eta)}(\mathbb{C}_d^{(0)})) \\
&\leqslant \mathrm{Const}(\mathscr{Z}_{j/2^n}^{(\eta)}(\mathbb{C}_d^{(0)}))\eta^2(1 + \eta^{\beta_1} 2^{-n}\beta_1/(\beta_1+1))^{-1/\beta_1}(1 + \eta^{\beta_2} 2^{-n}\beta_2/(\beta_2+1))^{-1/\beta_2} \\
&\quad \times \mathbb{P}_0\{|w(j\cdot 2^{-n}) - w(k\cdot 2^{-n})| > g_\kappa^{(2)}((j-k)\cdot 2^{-n})\},
\end{aligned} \tag{5.3}$$

where $\mathbb{P}_0$ is the Wiener measure on the space of continuous trajectories that start from the origin, the initial distribution m_η is give by (4.1), *and the historical aggregated mass $\mathscr{K}^{(\eta)}_{n,j,k}(t)$ is defined below formula* (2.3′).

Remark 5.3. Note that the constant on the right-hand side of (5.3) is in fact proportional to $\mathscr{Z}^{(\eta)}_{j/2^n}(\mathbb{C}^{(0)}_d)$, which is finite by the non-explosion property of $(2, d, \beta_1, \beta_2)$-historical processes.

Proof of Lemma 5.2. Straightforward and is based on the clearly evident facts that an individual superparticle alive at time $j \cdot 2^{-n}$ will survive up to time $t \geqslant (j+1)2^{-n}$ with probability

$$(1 + t \cdot \eta^{\beta_2} \beta_2/(\beta_2 + 1))^{-1/\beta_2} \leqslant (1 + 2^{-n} \eta^{\beta_2} \beta_2/(\beta_2 + 1))^{-1/\beta_2},$$

and an individual particle alive at time $j \cdot 2^{-n}$ will survive up to time $t \geqslant (j+1) \cdot 2^{-n}$, provided that it belongs to a superparticle alive at time t, with probability

$$(1 + t \cdot \eta^{\beta_1} \beta_1/(\beta_1 + 1))^{-1/\beta_1} \leqslant (1 + 2^{-n} \eta^{\beta_1} \beta_1/(\beta_1 + 1))^{-1/\beta_1}$$

(compare to (4.3)).

The remainder of the proof follows by noting that there are $\mathscr{Z}^{(\eta)}_{j/2^n}(\mathbb{C}^{(0)}_d)\eta^2$ particles at time $j/2^n$, the probability that any one of these particles follows a "bad" trajectory is equal to

$$\mathbb{P}_0\{|w(j \cdot 2^{-n}) - w(k \cdot 2^{-n})| > g^{(2)}_\kappa((j-k) \cdot 2^{-n})\},$$

and finally the probability that such a particle survives up to time $t \geqslant 2^{-n}$ is less than or equal to

$$(1 + \eta^{\beta_1} \cdot 2^{-n} \cdot \beta_1/(\beta_1 + 1))^{-1/\beta_1} \cdot (1 + \eta^{\beta_2} \cdot 2^{-n} \cdot \beta_2/(\beta_2 + 1))^{-1/\beta_2}. \qquad \square$$

Corollary 5.4. *For any $m \in M_F(M_F(\mathbb{R}^d))$, there exists a finite positive constant $C = C(m, d, \beta_1, \beta_2, \mathscr{Z}^{(\eta)}_{j/2^n}(\mathbb{C}^{(0)}_d))$ such that for any fixed real $t \geqslant 2^{-n}$,*

$$\begin{aligned}
&\mathscr{Q}_{0,m}(\mathscr{K}^{(\eta)}_{n,j,k}(t) > 0 \mid \mathscr{Z}^{(\eta)}_{j/2^n}(\mathbb{C}^{(0)}_d)) \\
&\leqslant C \cdot \eta^2 \cdot (1 + 2^{-n} \eta^{\beta_2} \beta_2/(\beta_2 + 1))^{-1/\beta_2} \cdot (1 + 2^{-n} \eta^{\beta_1} \beta_1/(\beta_1 + 1))^{-1/\beta_1} \\
&\quad \times \mathbb{P}_0\{|w(j \cdot 2^{-n}) - w(k \cdot 2^{-n})| > g^{(2)}_\kappa((j-k)2^{-n})\}.
\end{aligned} \tag{5.4}$$

Proof. Relies on an approximation of $m \in M_F(M_F(\mathbb{R}^d))$ by linear combinations of δ_δ-measures and a subsequent application of Lemma 5.2. $\square$

The next lemma is similar to [DV, estimate (2.3)].

Lemma 5.5. *For any* $\kappa > [d + 8 + 4/\beta_1 + 4/\beta_2]/[2(1 + 1/\beta_1 + 1/\beta_2)]$,

$$\mathcal{Q}_{0,m}\left\{\exists\{n_l\}\uparrow\infty \text{ such that } \exists 0 \leqslant k < j \leqslant 2^{n_l}, j-k \leqslant n_l^2\right.$$
$$\left.\text{such that } \sup_{t\geqslant 0} \mathcal{K}_{n_l,j,k}(2^{-n_l}+t) > 0\right\} = 0. \tag{5.5}$$

Proof. First note that by the extinction property of the $(2, d, \beta_1, \beta_2)$-historical process, we have

$$\lim_{L\to\infty} \mathcal{Q}_{0,m}\left(\sup_{t\geqslant 0} \mathcal{Z}_t(\mathbb{C}_d^{(0)}) \leqslant L\right) = 1,$$

and therefore it suffices to prove that

$$\mathcal{Q}_{0,m}\left\{\left(\exists\{n_l\}\uparrow\infty \text{ such that } \exists 0 \leqslant k < j \leqslant 2^{n_l}, j-k \leqslant n_l^2 \text{ such that}\right.\right.$$
$$\left.\left.\sup_{0\leqslant t\leqslant T} \mathcal{K}_{n_l,j,k}(2^{-n_l}+t) > 0\right) \cap \left(\sup_{t\geqslant 0} \mathcal{Z}_t(\mathbb{C}_d^{(0)}) \leqslant L\right)\right\} = 0 \tag{5.5'}$$

for $0 < L, T < \infty$. Now, fix $n(=n_l) \in \mathbb{N}$ and use the fact that

$$\left\{\sup_{0\leqslant t\leqslant T} \mathcal{K}_{n_l,j,k}(2^{-n_l}+t) > 0\right\}$$

is an open subset to obtain that

$$\mathcal{Q}_{0,m}\left\{\left(\exists 0 \leqslant k < j \leqslant 2^n, j-k \leqslant n^2 \text{ such that } \sup_{0\leqslant t\leqslant T} \mathcal{K}_{n,j,k}(2^{-n}+t) > 0\right)\right.$$
$$\left.\cap\left(\sup_{t\geqslant 0} \mathcal{Z}_t(\mathbb{C}_d^{(0)}) \leqslant L\right)\right\}$$
$$\leqslant \liminf_{\eta\to\infty} \mathcal{Q}_{0,m}^{(\eta)}\left\{\left(\exists 0 \leqslant k < j \leqslant 2^n, j-k \leqslant n^2 \text{ such that}\right.\right.$$
$$\left.\left.\sup_{0\leqslant t\leqslant T} \mathcal{K}_{n,j,k}^{(\eta)}(2^{-n}+t) > 0\right) \cap \left(\sup_{t\geqslant 0} \mathcal{Z}_t^{(\eta)}(\mathbb{C}_d^{(0)}) \leqslant L+1\right)\right\}$$
$$\leqslant \liminf_{\eta\to\infty} \mathcal{Q}_{0,m}^{(\eta)}\left\{\left(\max_{\substack{0\leqslant k<j\leqslant 2^n:\\ j-k\leqslant n^2}} \mathcal{K}_{n,j,k}^{(\eta)}(2^{-n}) > 0\right) \cap \left(\sup_{t\geqslant 0} \mathcal{Z}_t^{(\eta)}(\mathbb{C}_d^{(0)}) \leqslant L+1\right)\right\}$$
$$\leqslant \text{Const}\cdot 2^n \liminf_{\eta\to\infty}\left(\sum_{1\leqslant j\leqslant n^2} \mathcal{Q}_{0,m}^{(\eta)}\left\{\left(\mathcal{K}_{n,j,0}^{(\eta)}(2^{-n}) > 0\right)\right.\right.$$
$$\left.\left.\cap\left(\sup_{t\geqslant 0} \mathcal{Z}_t^{(\eta)}(\mathbb{C}_d^{(0)}) \leqslant L+1\right)\right\}\right). \tag{5.5''}$$

Estimating the above probabilities by the use of Corollary 5.4 with $t = 2^{-n}$, we obtain that the latter expression does not exceed

$$\lim_{\eta\to\infty}\Bigg(C(m,d,\beta_1,\beta_2,L)2^n\eta^2(1+2^{-n}\eta^{\beta_2}\beta_2/(\beta_2+1))^{-1/\beta_2}$$
$$\times(1+2^{-n}\eta^{\beta_1}\beta_1/(\beta_1+1))^{-1/\beta_1}$$
$$\times\sum_{1\leqslant j\leqslant n^2}\mathbb{P}_0\{|w(j\cdot 2^{-n})|>g_\kappa^{(2)}(j\cdot 2^{-n})\}\Bigg)$$
$$\leqslant \lim_{\eta\to\infty}\Bigg\{C(m,d,\beta_1,\beta_2,L)(2^n\cdot\eta^{-\beta_2}+\beta_2/(\beta_2+1))^{-1/\beta_2}$$
$$\times(2^n\cdot\eta^{-\beta_1}+\beta_1/(\beta_1+1))^{-1/\beta_1}$$
$$\times 2^{n(1+1/\beta_1+1/\beta_2)}\sum_{1\leqslant j\leqslant n^2}\mathbb{P}_0\{|w(1)|>(2^n/j)^{1/2}g_\kappa^{(2)}(j\cdot 2^{-n})\}\Bigg\},$$

where we have used the self-similarity property of the Wiener process.

It is relatively easy to see that the value of the above limit as $\eta\to\infty$ does not exceed

$$C_1(m,d,\beta_1,\beta_2,L)\cdot 2^{n(1+1/\beta_1+1/\beta_2)}\sum_{1\leqslant j\leqslant n^2}\mathbb{P}_0\{|w(1)|>(2^n/j)^{1/2}g_\kappa^{(2)}(j\cdot 2^{-n})\}. \tag{5.6}$$

Moreover, by [DHV3, formula (6)], the expression (5.6) and, therefore, the probability (5.5′) do not exceed

$$C_2(m,d,\beta_1,\beta_2,L)\cdot 2^{n(1+1/\beta_1+1/\beta_2)}\sum_{1\leqslant j\leqslant n^2}\exp\{-(2^n/j)\cdot(g_\kappa^{(2)}(j\cdot 2^{-n}))^2/2\}$$
$$\cdot((2^n/j)^{1/2}\cdot g_\kappa^{(2)}(j\cdot 2^{-n}))^{d-2}$$
$$\leqslant C_3(m,d,\beta_1,\beta_2,L)\cdot\sum_{1\leqslant j\leqslant n^2} j^{1+1/\beta_1+1/\beta_2}\left(\log\frac{2^n}{j}\right)^{-\kappa(1+1/\beta_1+1/\beta_2)+(d-2)/2}$$
$$\leqslant C_4(m,d,\beta_1,\beta_2,L)n^{2\cdot(2+1/\beta_1+1/\beta_2)}n^{-1-\kappa(1+1/\beta_1+1/\beta_2)-d/2}. \tag{5.7}$$

Recall that

$$\kappa>\frac{d+8+4/\beta_1+4/\beta_2}{2(1+1/\beta_1+1/\beta_2)}.$$

Hence, the expression on the right-hand side of (5.7) is the general term of a convergent series. Therefore, both expression (5.6) and the probability on the left-hand side of (5.5′) are also the general terms of a convergent series. A subsequent application of Borel–Cantelli arguments yields the assertion of the lemma. □

Proof of Theorem 2.5(ii). Develops along the same lines as that of [DV, Theorem 1.2]. The required estimate for the maximum over the grid is given by Lemma 5.5. We shall also use a modification of the standard technique (cf., e.g., McKean, 1969, p. 16) that involves an application of Lemma 5.5. However, we first reformulate relationship (2.6′) of Theorem 2.5(ii) in the following more convenient form. Specifically, relationship (2.6′) is equivalent to the statement that

$$\mathscr{X}_t((K_\kappa^{(2)}(\delta_2(\omega,\kappa),t))^c) = 0 \quad \forall t > 0, \quad \mathscr{Q}_{0,m}\text{-a.s.},$$

where function $K_\kappa^{(2)}$ is defined by (2.5).

Now, fix an arbitrary

$$\kappa > \kappa(\beta_1,\beta_2,d) = \frac{d + 8 + 4/\beta_1 + 4/\beta_2}{2(1 + 1/\beta_1 + 1/\beta_2)},$$

and recall that

$$\kappa_1 = (\kappa + \kappa(\beta_1,\beta_2,d))/2.$$

By Lemma 5.5, we know that for $\mathscr{Q}_{0,m}$-a.e. ω, $\mathscr{X}_t(\omega)$ is right continuous and there exists $n_1(\omega)$ such that for all $n \geqslant n_1(\omega)$,

$$\sup_{t\geqslant 0} \mathscr{X}_{j\cdot 2^{-n}+t}(\omega, \mathscr{G}^{(2)}_{n,j,k}) = 0.$$

In the remainder of the proof, we fix a pair $(\omega, n_1(\omega))$ for which this condition is satisfied. (Note that this is the value of n_1 that really matters, while ω enters only because it defines n_1.)

Now, by the right continuity of $\{\mathscr{X}_{t+s}\}_{s\geqslant 0}$ at zero, it suffices to show that

$$\mathscr{X}_{j\cdot 2^{-n}}(K_\kappa^{(2)}(\delta_2(\omega,\kappa), j\cdot 2^{-n})^c) = 0$$

for all pairs (j,n) with $n \geqslant n_1(\omega)$ and j such that $j\cdot 2^{-n} \geqslant t$. Hence, it suffices to find a $\delta_2(\omega,\kappa) > 0$ such that for all $n \geqslant n_1(\omega)$ and j with $j\cdot 2^{-n} \geqslant t$,

$$K_\kappa^{(2)}(\delta_2(\omega,\kappa), j\cdot 2^{-n}) \supset \bigcap\{(\mathscr{G}^{(2)}_{n,j,k})^c\colon 0 \leqslant k < j \leqslant 2^n,\ n \geqslant n_1(\omega)\}.$$

Let

$$\mathscr{W} \in \bigcap\{(\mathscr{G}^{(2)}_{n,j,k})^c\colon 0 \leqslant k < j \leqslant 2^n,\ n \geqslant n_1(\omega)\}.$$

Then for any integer $N = 2^n$ with $n \geqslant n_1(\omega)$,

$$\max_{\substack{0<k=j-i\leqslant(\log N)^2\\ 0\leqslant i<j\leqslant N}} \frac{|\mathscr{W}_{j/N} - \mathscr{W}_{i/N}|}{g_{\kappa_1}^{(2)}(k/N)} \leqslant 1. \tag{5.8}$$

Now, using the same arguments as those used between formulas (2.5) and (2.13) of [DV] or those used between formulas (9)–(16) of [DHV3], and applying Lemmas 5.1 and 5.5, one concludes that

$$|\mathscr{W}_t - \mathscr{W}_s| \leqslant g_\kappa^{(2)}(t-s)$$

for any $0 \leqslant t - s \leqslant \delta_2(\omega, \kappa)$, provided that $\delta_2(\omega, \kappa)$ is chosen so that

$$n(\delta_2(\omega,\kappa)) \geqslant \max(n_1(\omega), n(g_\kappa^{(2)}, C_5(\beta_1,\beta_2,d,L,\kappa_1))).$$

Therefore, $\mathscr{W} \in K_\kappa^{(2)}(\delta_2(\omega,\kappa),t)$. $\square$

Now, we proceed with the

Proof of Theorem 2.7(ii). The proof is similar to that of Theorem 2.5(ii); it involves a modification of the standard technique (cf., e.g., McKean, 1969, p. 16) and the use of arguments similar to those used in the proof of Theorem 2.5(ii) and Lemma 5.5.

Fix

$$\kappa > \hat{\kappa}(\beta_1,\beta_2,d) = \frac{d/2+2}{1/\beta_1 + 1/\beta_2} + 2$$

and recall that

$$\kappa_2 := (\kappa + \hat{\kappa}(\beta_1,\beta_2,d))/2.$$

By analogy with the proof of Lemma 5.5, it suffices to find a $\delta_2^*(\omega,\kappa) > 0$ such that for all $n \geqslant n_1(\omega)$ and j with $j \cdot 2^{-n} \geqslant t$,

$$\hat{K}_\kappa^{(2)}(\delta_2^*(\omega,\kappa), j\cdot 2^{-n}) \supset \bigcap\{(\mathscr{G}_{n,j,k}^{(2)})^c : 0 \leqslant k < j \leqslant 2^n,\ n \geqslant n_1(\omega)\},$$

where $\hat{K}_\kappa^{(2)}$ is defined by (2.8), and $\hat{\mathscr{G}}_{n,j,k}^{(2)}$ is defined below (2.4′). To this end, choose a $\hat{\mathscr{W}}$ belonging to

$$\bigcap\{(\hat{\mathscr{G}}_{n,j,k}^{(2)})^c : 0 \leqslant k < j \leqslant 2^n,\ n \geqslant n_1(\omega)\}$$

and show that such a $\hat{\mathscr{W}}$ also belongs to $\hat{K}_\kappa^{(2)}(\delta_2^*(\omega,\kappa), j\cdot 2^{-n})$.

At this stage, we introduce

$$\hat{\mathscr{K}}_{n,j,k}(t) := \mathscr{X}_{j\cdot 2^{-n}+t}(\hat{\mathscr{G}}_{n,j,k}^{(2)}) \quad \text{and} \quad \hat{\mathscr{K}}_{n,j,k}^{(\eta)}(t) := \mathscr{X}_{j\cdot 2^{-n}+t}^{(\eta)}(\hat{\mathscr{G}}_{n,j,k}^{(2)}),$$

where $\hat{\mathscr{G}}_{n,j,k}^{(2)}$ is defined below formula (2.4′) (compare to notation introduced below formula (2.3′)). Now, fix $n \in \mathbb{N}$ and $0\cdot \leqslant k \leqslant 2^n$, and estimate the following expression:

$$2^n \mathscr{Q}_{0,m}\left\{\left(\exists j\colon k \leqslant j \leqslant k+n^2 \text{ such that } \hat{\mathscr{K}}_{n,j,k}(2^{-n}) > 0\right) \cap \left(\sup_{t \geqslant 0} \mathscr{X}_t(\mathbb{C}_d^{(0)}) \leqslant L\right)\right\}, \tag{5.9}$$

where $L > 0$ is fixed. By analogy with the arguments given below formula (5.5′) which are used in the proof of Lemma 5.5, and in view of the BPS-2 approximation, this expression does not exceed

$$\begin{aligned}
&\lim_{\eta\to\infty} \mathscr{Q}_{0,m}^{(\eta)}\left\{\left(\exists j\colon k \leqslant j \leqslant k+n^2 \text{ such that } \hat{\mathscr{K}}_{n,j,k}^{(\eta)}(2^{-n}) > 0\right) \cap \left(\sup_{t\geqslant 0} \mathscr{X}_t^{(\eta)}(\mathbb{C}_d^{(0)}) \leqslant L\right)\right\} \\
&\quad \leqslant \lim_{\eta\to\infty} \mathscr{Q}_{0,m}^{(\eta)}\left\{\left(\max_{k\leqslant j\leqslant k+n^2} \hat{\mathscr{K}}_{n,j,k}^{(\eta)}(2^{-n}) > 0\right) \cap \left(\sup_{t\geqslant 0} \mathscr{X}_t^{(\eta)}(\mathbb{C}_d^{(0)}) \leqslant L\right)\right\}.
\end{aligned}$$

We estimate the conditional probabilities by analogy with Lemma 5.5 with $t = 2^{-n}$ and with arguments given in the proof of Theorem 1 of [DHV3, formulas (3)–(9)] to obtain that the above limit does not exceed

$$C_1(m,d,\beta_1,\beta_2,L)\lim_{\eta\to\infty}\Bigg(\eta^2(1+2^{-n}\eta^{\beta_2}\beta_2/(\beta_2+1))^{-1/\beta_2}$$

$$\times(1+2^{-n}\eta^{\beta_1}\beta_1/(\beta_1+1)^{-1/\beta_1}\,\mathbb{P}_0\left\{\max_{1\leqslant j\leqslant n^2}|w(j\cdot 2^{-n})|/h_\kappa^{(2)}(j\cdot 2^{-n})>1\right\}\Bigg)$$

$$\leqslant C_2(m,d,\beta_1,\beta_2,L)\cdot 2^{\,n(1/\beta_1+1/\beta_2)}\,n^2\,\mathbb{P}_0\left\{|w(1)|>(2^n/n^2)^{1/2}h_\kappa^{(2)}(n^2\cdot 2^{-n})\right\}.$$

Now, it is relatively easy to see that the above expression does not exceed

$$C_3(m,d,\beta_1,\beta_2,L)\,2^{n(1/\beta_1+1/\beta_2)}\,n^2\,\mathbb{P}_0\{|w(1)|>(2^n/n^2)^{1/2}h_\kappa^{(2)}(n^2\cdot 2^{-n})\}$$

$$\leqslant C_4(m,d,\beta_1,\beta_2,L)\,2^{n(1/\beta_1+1/\beta_2)}\,n^2$$

$$\exp\{-2^n/n^2)\cdot(h_\kappa^{(2)}(n^2\cdot 2^{-n}))^2/2\}\cdot((2^n/n^2)^{1/2}h_\kappa^{(2)}(n^2\cdot 2^{-n}))^{d-2}$$

$$\leqslant C_5(m,d,\beta_1,\beta_2,L)\,n^{2+2\cdot(1/\beta_1+1/\beta_2)}\left(\log\frac{2^n}{n^2}\right)^{-\kappa\cdot(1/\beta_1+1/\beta_2)+(d-2)/2}$$

$$\leqslant C_6(m,d,\beta_1,\beta_2,L,\kappa)\,n^{-1-\kappa(1/\beta_1+1/\beta_2)+d/2+2+2(1/\beta_1+1/\beta_2)}. \tag{5.10}$$

Since

$$\kappa>\frac{d/2+2}{1/\beta_1+1/\beta_2}+2,$$

the expression on the right-hand side of (5.10) is the general term of a convergent series. Hence, the expression (5.9) is also a general term of a convergent series. Then an argument analogous to that given in the proof of Theorem 1 of [DHV3] yields the following estimate for the grid:

$$\max_{0<i\leqslant(\log N)^2}\frac{|\hat{\mathscr{W}}_{t+i/N}-\hat{\mathscr{W}}_t|}{h_{\kappa_2}^{(2)}(i/N)}\leqslant 1. \tag{5.11}$$

for any integer $N=2^n$ with $n\geqslant n_1$.

Subsequently, we apply the same arguments as those used in [DV, Section 2 (see pp. 242–243 therein] to derive that (5.11) implies

$$|\hat{\mathscr{W}}_{t+u}-\hat{\mathscr{W}}_t|\leqslant h_{\kappa_2}^{(2)}(u)+C_3(\beta_1,\beta_2,d,L,\kappa_2)h_{\kappa_2}^{(2)}(1/2^n). \tag{5.12}$$

Finally, a combination of (5.12) with Lemma 5.1(ii) implies that

$$\hat{\mathscr{W}}\in\hat{K}_\kappa^{(2)}(\delta_2^*(\omega,\kappa),j\cdot 2^{-n}).\quad\square$$

We now proceed with the

Proof of Theorem 2.1(iii). In order to get an upper estimate, we apply relationship (2.10′) of Corollary 2.8(ii) from Theorem 2.7(ii) with $t = 0$ and $m = \delta_{\delta_0}$. In particular, we get that

$$S(Z_0) = \{0\}, \quad \mathscr{P}_{\delta_{\delta_0}}\text{-a.s.}$$

Hence,

$$S(Z_0)^{h_\kappa^{(2)}(s)} = \overline{\mathbb{B}(0, h_\kappa^{(2)}(s))}, \quad \mathscr{P}_{\delta_{\delta_0}}\text{-a.s.}$$

A subsequent application of Corollary 2.8(ii) yields that for each

$$\kappa > \frac{d/2 + 2}{1/\beta_1 + 1/\beta_2} + 2$$

and for each $0 < s < \hat{\delta}_*^{(2)}(\omega, \kappa)$,

$$\mathscr{P}_{\delta_{\delta_0}}\{S(Z_u) \subseteq \overline{\mathbb{B}(0, h_\kappa^{(2)}(s))} \text{ for } 0 \leqslant u \leqslant s\} = 1. \tag{5.13}$$

In particular, a combination of formula (2.4′) for $h_\kappa^{(2)}(\cdot)$ and (5.13) implies that for any positive $\kappa > \dfrac{d/2 + 2}{1/\beta_1 + 1/\beta_2} + 2$,

$$\mathscr{P}_{\delta_{\delta_0}}\left\{ \sup_{0 \leqslant u \leqslant t} r^{(2)}(u) \leqslant \sqrt{2\left(\frac{1}{\beta_1} + \frac{1}{\beta_2}\right) t \left(\log\frac{1}{t} + \kappa \log\log\frac{1}{t}\right)} \right.$$
$$\left. \text{for all sufficiently small positive } t \right\} = 1. \tag{5.14}$$

Combining (5.14) with (4.10), we obtain the result that for each sufficiently small positive ε,

$$\mathscr{P}_{\delta_{\delta_0}}\left\{ \sqrt{2\left(\frac{1}{\beta_1} + \frac{1}{\beta_2}\right) t \left(\log\frac{1}{t} + \left(\frac{d-2}{2(1/\beta_1 + 1/\beta_2)} - \varepsilon\right) \log\log\frac{1}{t}\right)} \right.$$
$$\leqslant \sup_{0 \leqslant u \leqslant t} r^{(2)}(u) \leqslant \sqrt{2\left(\frac{1}{\beta_1} + \frac{1}{\beta_2}\right) t \left(\log\frac{1}{t} + \left(\frac{d/2 + 2}{1/\beta_1 + 1/\beta_2} + 2 + \varepsilon\right) \log\log\frac{1}{t}\right)}$$
$$\left. \text{for all sufficiently small positive } t \right\} = 1.$$

Dividing all three items of the inequality under the probability sign by $\sqrt{2(1/\beta_1 + 1/\beta_2) t \log\frac{1}{t}}$ and keeping in mind that $\sqrt{1 + \theta} = 1 + \theta/2 + O(\theta^2)$ as $\theta \to 0$, we immediately obtain (2.1″). □

Proof of Theorem 2.1(ii). Follows along the same lines as the proof of Theorem 2.1(iii) and therefore is omitted.

6. Application to the Hausdorff dimension of the closed support of the aggregated process

It turns out that arguments similar to those used in the proof of Lemma 5.6 of the previous section, along with slight variants of arguments used in the proof of Lemma 9.3.3.7 of [D], imply the following almost-sure upper estimate for the Hausdorff dimension of the closed support of the aggregated process:

Proposition 6.1. *The closed support $S(Z_t)$ of the aggregated process Z_t satisfies*

$$\mathscr{P}_{\delta_{\delta_0}}\left\{\dim S(Z_t) \leqslant 2\left(\frac{1}{\beta_1}+\frac{1}{\beta_2}\right) \quad \forall t>0\right\}=1. \tag{6.1}$$

Remark 6.2. Let us emphasize that a combination of (6.1) with Theorem 4.3.1 of Wu (1992) suggests that for $d \geqslant 2(1/\beta_1 + 1/\beta_2)$, the Hausdorff dimension of the closed support of the aggregated process Z_t associated with the super-2 process X_t is in fact *equal to* $2(1/\beta_1 + 1/\beta_2)$. Indeed, in the special case where $d \geqslant 4$, $X_0 = \delta_{\delta_x}\,dx$, and $\beta_1 = \beta_2 = 1$, Wu (1992, Theorem 4.3.1) established that the Hausdorff dimension of $S(Z_t)$ is *greater than or equal to four.* However, Wu used arguments based on an application of Zähle's lemma (cf. Zähle (1988)) involving second moment measures in order to derive this result, and that technique is not applicable in the general case.

Proof of Proposition 6.1. Is in the same spirit as those of Theorem 9.3.3.5 and Lemma 9.3.3.7 of [D]. We first give some heuristic ideas behind our arguments.

Let $N(s,t)$ denote the number of level-1 clusters starting at time s whose lifetimes are greater than t. Then, denote the number of level-1 clusters at instant $j/2^n$ of age 2^{-n} by

$$N_n(j) := N((j-1)\cdot 2^{-n}, j\cdot 2^{-n}).$$

Note that conditioned on the total mass $Z_t(\mathbb{C}_d^{(0)})$ of the aggregated process having value L at time $t = (j-1)2^{-n}$, the distribution of $N_n(j)$ is compound Poisson, with Laplace transform given by

$$\exp\left\{L(\beta_2/(\beta_2+1))^{-1/\beta_2}2^{n/\beta_2}\left\{\exp(\beta_1/(\beta_1+1))^{-1/\beta_1}2^{n/\beta_1}(e^{-s}-1)-1\right\}\right\}$$

(compare to the Laplace transform given in Corollary 4.2, that differs only by conditioning on non-extinction).

In particular, the above representation implies that

$$\mathscr{E}_{\delta_{\delta_0}} N_n(j) = L(\beta_2/(\beta_2+1))^{-1/\beta_2}(\beta_1/(\beta_1+1))^{-1/\beta_1}2^{n(1/\beta_1+1/\beta_2)},$$

where $\mathscr{E}_{\delta_{\delta_0}}$ denotes the expectation with respect to $\mathscr{P}_{\delta_{\delta_0}}$.

Intuitively, each level-1 cluster of age t represents a subpopulation alive at time $s+t$ and having a single common ancestor located at a point x at time s. We will obtain a covering of the closed support $S(Z_t)$ of Z_t in the time interval $[(j-1)2^{-n}, j\cdot 2^{-n}]$ by closed balls of radius r_n by decomposing the random measure

Z_t into a finite number $N_n(j)$ of clusters starting at time $(j-1)2^{-n}$ at points $\{x_l\}_{l=1,\ldots,N_n(j)}$. Subsequently, we can use arguments similar to those given in the proof of Theorem 2.7(ii) for the derivation of an upper estimate for the probability of the event that the closed ball $\overline{\mathbb{B}(x_l,r_n)}$ covers a particular evolving level-1 cluster during the time interval $[(j-1)2^{-n}, j\cdot 2^{-n}]$.

We now apply Lemma 9.3.3.7.ii of [D] to establish that the random variable $N_n(j)$ is in fact of the same order as its expectation. Also, note that Lemma 9.3.3.7.ii of [D] was obtained by the use of purely probabilistic methods, namely, by employing the exponential Chebyshev inequality and Borel–Cantelli arguments.

Lemma 6.3. *Let L and L_* be two arbitrary positive constants. Let $\tau_{L,L_*} := \min(1, \inf \{t\colon \{X_t(M_F(M_F(\mathbb{R}^d))) > L\} \cup \{Z_t(\{\mu\colon \mu(\mathbb{R}^d) > L_*\}) > 0\}\})$. Then there exists $n_0(\omega)$ such that for all $n \geqslant n_0(\omega)$,*

$$\max_{1 \leqslant j \leqslant 2^n \tau_{L,L_*}} N_n(j) \leqslant C(L, L_*, \beta_1, \beta_2)\, 2^{n(1/\beta_1 + 1/\beta_2)}, \quad \mathscr{P}_{\delta_{\delta_0}}\text{-a.s.}$$

Proof. Indeed, Lemma 9.3.3.7.ii of [D] (related to the $(2,d,\beta)$-super-1 process U_t) implies that

$$\max_{1 \leqslant j \leqslant 2^n \tau_{L,L_*}} N_n^*(j) \leqslant C(L_*, \beta)\, 2^{n/\beta}, \quad \mathbf{P}_{m\cdot\delta_0}\text{-a.s.},$$

where $N_n^*(j)$ denotes the number of (*level*-1) clusters at instant $j/2^n$ of age 2^{-n} of the $(2,d,\beta)$-super-1 process $\tilde{U}_t$ with state space $M_F(\mathbb{R}^d)$ (considered above Theorem 2.5), provided that $\tilde{U}_t(\mathbb{R}^d) \leqslant L_*$. We apply this result (with $\beta = \beta_2$) to obtain that the number of living *level*-2 clusters of age 2^{-n} at time instant $j/2^n$ does not exceed $C_1(L,\beta_2)\cdot 2^{n/\beta_2}$, $\mathscr{P}_{\delta_{\delta_0}}$-a.s. Recall that each level-2 cluster can be viewed as a Poisson number of level-1 clusters. We successively apply Lemma 9.3.3.7.ii of [D] to conclude that the number of level-1 clusters in each two-level cluster does not exceed

$$C_2(L_*, \beta_1)\cdot 2^{n/\beta_1}, \quad \mathscr{P}_{\delta_{\delta_0}}\text{-a.s.},$$

which easily yields the assertion of the lemma. □

We now complete the proof of Proposition 6.1. We first split the time interval $[0, \tau_{L,L_*}]$ into intervals $[(j-1)\cdot 2^{-n}, j\cdot 2^{-n}]$ of length 2^{-n} $(j = 1, \ldots, 2^n[\tau_{L,L_*}])$. We then cover by balls of radius $r_n = C_1 \cdot 2^{-n/2} n^{1/2}$ each of the $N_n(j)$ level-1 cluster birth points at time $(j-1)2^{-n}$ that has surviving descendants at time $j\cdot 2^{-n}$. Recall that according to Lemma 6.3, for each $j\cdot 2^{-n} \leqslant \tau_{L,L_*}$, the number of such level-1 cluster birth points does not exceed $C(L, L_*, \beta_1, \beta_2)\cdot 2^{n(1/\beta_1 + 1/\beta_2)}$. Now, set

$$\Xi_n(j) := \bigcup_{l=1}^{N_n(j)} \overline{\mathbb{B}(x_l, r_n)}.$$

Applying arguments given in the proof of Theorem 2.7(ii) to each of these single-level clusters, we obtain that

$$\begin{aligned}
&\mathscr{P}_{\delta_{\delta_0}}\{Z_u(\overline{\Xi_n(j)^c}) > 0 \text{ for some } 0 \leqslant u \leqslant t,\\
&\forall t \in [(j-1)2^{-n}, j\cdot 2^{-n}] \quad \forall j/2^n \leqslant \tau_{L,L_*}\}\\
&\qquad \leqslant C(L, L_*, \beta_1, \beta_2)\cdot 2^{n(1+1/\beta_1+1/\beta_2)} n^{(d-2)/2} \exp\left\{-C_1^2\cdot 2^n\cdot 2^{-n}\cdot\frac{n}{2}\right\}.
\end{aligned}$$

The latter expression is the general term of a convergent series for sufficiently large C_1. Therefore,

$$\begin{aligned}
&\mathscr{P}_{\delta_{\delta_0}}\{S(Z_t) \subset \Xi_n(j) \quad \forall t \in [(j-1)\cdot 2^{-n}, j\cdot 2^{-n}\} \text{ for each } j/2^n \leqslant \tau_{L,L_*},\\
&\qquad j \in \mathbb{N}, \text{ for all sufficiently large } n\} = 1.
\end{aligned}$$

Now, set

$$\phi_{\beta_1,\beta_2}(x) := x^{2/\beta_1 + 2/\beta_2}(\log(1/x))^{-1/\beta_1 - 1/\beta_2},$$

and denote the Hausdorff measure corresponding to this function by ϕ_{β_1,β_2}-$m(\cdot)$. Recall that for a continuous, strictly increasing function ϕ on $[0,\infty)$ with $\varphi(0) = 0$, the *Hausdorff φ-measure* of a set $A \in \mathbb{R}^d$ is defined as

$$\varphi\text{-}m(A) := \lim_{\delta\downarrow 0}(\varphi)_\delta\text{-}m(A),$$

where

$$(\varphi)_\delta\text{-}m(A) := \inf\left\{\sum_k \varphi(r_k) : \bigcup_k \overline{\mathbb{B}(x_k, d_k/2)} \supseteq A,\ d_k < \delta\right\}.$$

Here $\{\overline{\mathbb{B}(x_k, d_k/2)}\}$ denotes the closed ball centered at x_k with radius $d_k/2$.

Employing this notation, and in view of the above inequality for the probability distribution of $Z_u(\overline{\Xi_n(j)^c})$, we conclude that for all sufficiently large n,

$$\begin{aligned}
\varphi_{\beta_1,\beta_2}\text{-}m(S(Z_t)) &\leqslant \lim_{n\to\infty}(\varphi_{\beta_1,\beta_2})_{2r_n}\text{-}m(S(Z_t)) \leqslant \lim_{n\to\infty}(\varphi_{\beta_1,\beta_2})_{2r_n}\text{-}m(\Xi_n(j))\\
&\leqslant \lim_{n\to\infty}(C(L, L_*, \beta_1, \beta_2)\cdot 2^{n(1/\beta_1+1/\beta_2)}\cdot\varphi_{\beta_1,\beta_2}(2r_n)) < \infty. \quad \square
\end{aligned}$$

Acknowledgement

We thank Ed Perkins for his advice during the preparation of the preliminary version of this work and Jean-François Le Gall for pointing out the argument that lead to Remark 2.2(ii). We also thank an anonymous referee for a careful reading and very useful suggestions.

References

M.T. Barlow and E.A. Perkins, On the filtration of historical Brownian motion, Ann. Prob. 22 (1994) 1273–1294.

T. Bojdecki and L.G. Gorostiza, Self-intersection local time for Gaussian $S'(\mathbf{R}^d)$-processes: existence, path continuity and examples, Tech. Rept No. 262, LRSP, Carleton Univ. (Ottawa, Canada, 1995).

K.L. Chung, P. Erdös and T. Sirao, On the Lipschitz's condition for Brownian motions, J. Math. Soc. Japan 11 (1959) 263–274.

M. Csörgő and P. Révész, Strong Approximation in Probability and Statistics (Academic Press, New York, 1981).

D.A. Dawson [D], Measure-valued Markov processes, in: P.L. Hennequin, ed., Ecole d'Eté de Probabilités de Saint-Flour XXI – 1991, Lecture Notes in Mathematics, No. 1451 (Springer, Berlin, 1993) pp. 191–260.

D.A. Dawson and K.J. Hochberg, A multilevel branching model, Adv. Appl. Probab. 23 (1991) 701–715.

D.A. Dawson, K.J. Hochberg and V. Vinogradov [DHV1], On path properties of super-2 processes I, in: D.A. Dawson, ed., Measure-valued Processes, Stochastic Partial Differential Equations and Interacting Systems, CRM Proceedings and Lecture Notes No. 5 (AMS, Providence, 1994) pp. 69–82.

D.A. Dawson, K.J. Hochberg and V. Vinogradov [DHV2], On path properties of super-2 processes II, in: M.G. Cranston and M.A. Pinsky, eds., Proc. Symp. in Pure Mathematics Series Vol. 57 (AMS, Providence, 1995), pp. 385–403.

D.A. Dawson, K.J. Hochberg and V. Vinogradov [DHV3], The Wiener process revisited, in: Yu. V. Prohorov, A.N. Shiryaev, S. Watanabe, M. Fukushima, eds., Probability Theory and Mathematical Statistics. Proc. 7th Japan–Russia Symp. (World Scientific, Singapore, 1996a, in press).

D.A. Dawson, K.J. Hochberg and V. Vinogradov, On weak convergence of branching particle systems undergoing spatial motion, in: J.-P. Fouque, K.J. Hochberg, and E. Merzbach, eds., Stochastic Analysis: Random Fields and Measure-Valued Processes, Israel Math. Conf. Proc., Vol. 10 (AMS, Providence, 1996b), pp. 65–79.

D.A. Dawson, K.J. Hochberg and Y. Wu, Multilevel branching systems, in: T. Hida, H.H. Kuo, J. Potthoff and L. Streit, eds., White Noise Analysis: Mathematics and Applications (World Scientific, Singapore, 1990) pp. 93–107.

D.A. Dawson and E.A. Perkins, Historical Processes, Memoirs of the American Mathematical Society No. 93 (AMS, Providence, 1991).

D.A. Dawson and V. Vinogradov [DV], Almost-sure path properties of $(2, d, \beta)$-superprocesses, Stochastic Proc. Appl. 51 (1994) 221–258.

D.A. Dawson and Y. Wu, Multilevel multitype branching models of an information system, in: K. Athreya and P. Jagers, eds., Proc. IMA Workshop on Branching Processes, IMA (1996, in press).

E.B. Dynkin, Superprocesses and partial differential equations, Ann. Probab. 21 (1993) 1185–1262.

N. El Karoui and S. Roelly, Propriétés de martingales, explosion et représentation de Lévy-Khintchine d'une classe de processus de branchement á valeurs mesures, Stoch. Proc. Appl. 38 (1991) 239–266.

L.G. Gorositza, K.J. Hochberg and A. Wakolbinger, Persistence of a critical super-2 process, J. Appl. Prob. 32 (1995) 534–540.

K.J. Hochberg, Hierarchically structured branching populations with spatial motion, Rocky Mountain J. Math. 25 (1995) 269–283.

K.J. Hochberg and A. Wakolbinger, Non-persistence of two-level branching particle systems in low dimensions. in: A. Etheridge, ed., Stochastic Partial Differential Equations, London Mathematical Society Lecture Note Series No. 216 (Cambridge Univ. Press, Cambridge, 1995) pp. 126–140.

H.P. McKean, Stochastic Integrals (Academic Press, New York, 1969).

C. Mueller and E.A. Perkins, The compact support property for solutions to the heat equation with noise, Probab. Theory Rel. Fields 93 (1992) 325–358.

R. Tribe, Path properties of superprocesses, Ph.D. Thesis, Univ. British Columbia (1989).

Y. Wu, Dynamic particle systems and multilevel measure branching processes, Ph.D. Thesis, Carleton Univ. (Ottawa, Canada, 1992).

Y. Wu, Multilevel birth and death particle system and its continuous diffusion, Adv. Appl. Probab. 25 (1993) 549–569.

Y. Wu, Asymptotic behavior of the two level measure branching process, Ann. Probab. 22 (1994a) 854–874.

Y. Wu, A three level particle system and existence of general multilevel measure-valued processes. in: D.A. Dawson, ed., Measure-valued Processes, Stochastic Partial Differential Equations and Interacting Systems, CRM Proceedings and Lecture Notes, No. 5 (AMS, Providence, 1994b) pp. 233–241.

U. Zähle, Self-similar random measures I. Notion, carrying Hausdorff dimension and hyperbolic distribution, Probab. Theory Rel. Fields 80 (1988) 79–100.

V.M. Zolotarev, More exact statements of several theorems in the theory of branching processes, Theory Probab. Appl. 2 (1957) 245–253.